

Evolution spezifizierender Proteine der Ordnung Brassicales

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades einer
Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von

Jennifer-Christin Kuchernig

aus Braunschweig

1. Referent: Professorin Dr. Ute Wittstock
 2. Referent: apl. Professor Dr. Robert Hänsch
- eingereicht am: 25.07.2012
- mündliche Prüfung (Disputation) am: 28.11.2012

Druckjahr 2013

VORVERÖFFENTLICHUNGEN DER DISSERTATION

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen:

Kuchernig, J.-C., Burow, M., Wittstock, U.: Evolution of specifier proteins in glucosinolate-containing plants. BMC Evolutionary Biology: accepted (2012).

Kuchernig, J.-C., Backenköhler, A., Lübbecke, M., Burow, M., Wittstock, U.: A thiocyanate-forming protein generates multiple products upon allylglucosinolate breakdown in *Thlaspi arvense*. Phytochemistry 72: 1699-1709 (2011).

Tagungsbeiträge:

Kuchernig, J.-C., Burow, M., Wittstock, U.: Evolution of specifier proteins in the Brassicales. (Poster). DPhG Jahrestagung, Braunschweig (2010).

Kuchernig, J.-C., Burow, M., Backenköhler, A., Wittstock, U.: Identification of specifier proteins from glucosinolate-containing plants. (Poster). The 2nd Conference on Glucosinolates, Helsingør (2009).

DANKSAGUNG

Meinen besonderen Dank möchte ich Frau Prof. Dr. Ute Wittstock für die Bereitstellung des interessanten und aktuellen Themas dieser Dissertation aussprechen. Ihr Vertrauen und Interesse am Fortgang meiner Arbeit und der Freiraum, eigene Ideen selbständig verwirklichen zu dürfen, sowie die ständige Unterstützung, haben meine Begeisterung für die Forschung stets aufrechterhalten.

Herrn Prof. Dr. Robert Hänsch danke ich herzlich für die freundliche Übernahme des Korreferates sowie weitere in jeder Hinsicht geleistete Motivation und Unterstützung.

Herrn Dr. Rainer Lindigkeit möchte ich für die Hilfe bei jeglichen technischen und IT-Problemen und der Betreuung der Praktika danken. Bei Herrn Dr. Till Beuerle bedanke ich mich für die Unterstützung und schnelle Lösung jeglicher analytischer Probleme.

Loretta Heise danke ich für die geduldige Anzucht und Pflege vieler Pflanzen. Claudine Theuring möchte ich für die Unterstützung sowie Rat und Tat in der Analytik danken. Einen Dank für die liebevolle Einarbeitung möchte ich Dr. Meike Burow aussprechen. Anita Backenköhler danke ich besonders für ihre Denkanstöße, die Weitergabe ihrer enormen Erfahrungen die zum Gelingen dieser Arbeit maßgeblich beigetragen haben und eine schöne Zeit. Ein ganz besonderer Praktikant war Tobias Wille, der mich mit seiner unermüdlichen Motivation und Freundschaft jederzeit tatkräftig unterstützt hat.

Weiterhin danke ich allen Mitgliedern der Arbeitsgruppe Wittstock für die ausgesprochen nette Arbeitsatmosphäre im Labor und im Büro. Im Besonderen gilt dieser Dank Antje Plock, Dr. Inis Winde, Dr. Rene Müller, Anja Losansky, Dr. Einar Stauber, Frauke Gumz, Beena M.R., Malte Büttner, Luise Cramer, Dr. Annika Reinhard, Dr. Leif Barleben, Anja Böhme, Maike van Ohlen, Marion Wiggermann und allen Diplomanden und Praktikanten während meiner Laborzeit. Dr. Maren Luetge, Cornelia Hüttner und Andreas Müller der Arbeitsgruppe Beerhus danke ich für ihre investierte Zeit, gute Ratschläge und wunderbare Gespräche zu jeder Gelegenheit.

Ich danke ganz besonders meinen Eltern, die mich immer ermutigt haben meinen eigenen Weg zu gehen und mich bei jedem Schritt unterstützt und begleitet haben.

Ein besonderer Dank gilt meiner Familie und Thorsten Wille für ihre Geduld und Unterstützung während dieser Zeit.

Für meine Eltern



INHALTSVERZEICHNIS

Inhaltsverzeichnis	I
Symbol- und Abkürzungsverzeichnis	III
1. Einleitung	1
1.1 Das Glucosinolat-Myrosinase-System als Beispiel für ein pflanzliches Abwehrsystem.	1
1.2 Glucosinolat-Biosynthese und ihr evolutiver Hintergrund	3
1.2.1 Schritt I: Kettenverlängerung der Ausgangsaminosäure	4
1.2.2 Schritt II: Bildung der Glucosinolatgrundstruktur	4
1.2.3 Schritt III: Sekundäre Modifikation der Glucosinolatseitenkette	6
1.3 Kompartimentierung	7
1.4 Myrosinasen und ihr evolutiver Hintergrund	9
1.5 Spezifizierende Proteine als Teil des Glucosinolat-Myrosinase-Systems	10
1.6 Funktionen freigesetzter Glucosinolat-Hydrolyseprodukte und ihre Rolle im „evolutiven Wettrennen“	14
1.7 Fragestellung und Zielsetzung der vorliegenden Arbeit	17
2. Ergebnisse	19
2.1 Phytochemisches Screening zur Auswahl geeigneter Glucosinolat-haltiger Spezies	19
2.2 Wahl einer Klonierungsstrategie	24
2.3 Identifizierung von cDNA Sequenzen putativer spezifizierender Proteine	27
2.4 Funktionelle Charakterisierung der identifizierten spezifizierenden Proteine	31
2.5 Datenbankrecherche zur Identifizierung putativer spezifizierender Proteine	34
2.6 Sequenzanalyse spezifizierender Proteine	35
2.6.1 Sequenzidentitäten	35
2.6.2 Strukturanalysen	36
2.7 Phylogenetische Analyse zur Aufklärung der evolutiven Abstammung spezifizierender Proteine	37
2.7.1 Alignment	37
2.7.2 Substitutionsmodelle	37
2.7.3 Generation phylogenetischer Stammbäume	39
2.8 Paarweise Berechnung von Aminosäuredistanzen und Nukleotiddistanzen spezifizierender Proteine	46
2.9 Berechnung der Anzahl synonyme und nichtsynonyme Substitutionen zur Bestimmung des Selektionsdruckes	49
2.10 Identifizierung spezifizierender Proteine in weiteren Spezies der Ordnung Brassicales	52
2.10.1 Anwendung einer PCR-Strategie	52
2.11 Charakterisierung des Thiocyanat-formenden Proteins aus <i>T. arvense</i>	59

3. Diskussion.....	63
3.1 Klonierung und Charakterisierung spezifizierender Proteine.....	63
3.2 Phylogenetische Analysen spezifizierender Proteine	70
3.2.1 Ist At3g07720 ein Vorläuferprotein spezifizierender Proteine?	73
3.2.2 Sind NSPs die Basis der Entwicklung zu ESPs und TFPs?	74
3.2.3 Sind TFPs zwei Mal unabhängig voneinander aus ESPs entstanden oder kam es zu einem multiplen Verlust?.....	77
3.3 Mechanismen der Evolution der spezifizierenden Proteine	79
3.4 Charakterisierung des TaTFP aus <i>T. arvense</i>	83
3.5 Ausblick	85
4. Zusammenfassung	89
5. Material und Methoden	90
5.1 Material.....	90
5.1.1 Pflanzenmaterial	90
5.1.2 Bakterienstämme und Plasmidvektoren	92
5.1.3 Primer.....	93
5.1.4 Enzyme	93
5.1.5 Kits	94
5.1.6 Weitere Reagenzien und Materialien	94
5.2 Methoden	95
5.2.1 Molekularbiologische Methoden.....	95
5.2.2 Proteinbiochemische Methoden.....	111
5.2.3 Chemisch-analytische Methoden	114
5.2.4 Bioinformatische Methoden	117
5.3 Medien.....	119
5.4 Puffer und Lösungen	120
6. Anhang	134
6.1 Alignments zum Design von degenerierten Primern	134
6.2 Glucosinolatprofile und -hydrolyseprodukte der untersuchten Spezies	136
6.3 Paarweiser Identitätsvergleich aller Sequenzen spezifizierender Proteine mittels ClustalW	142
6.4 Datenmatrizen der Berechnung von d_{aa} und d_{na}	143
6.5 Datenmatrizen der Berechnung von $\omega = d_N / d_S$	145
6.6 Glucosinolatprofil in <i>T. arvense</i>	147
6.7 Primer.....	148

SYMBOL- UND ABKÜRZUNGSVERZEICHNIS

Die in dieser Arbeit verwendeten chemischen Namen folgen den Regeln für die Verwendung von Abkürzungen und Symbolen der „International Union of Pure and Applied Chemistry“ (IUPAC) und der internationalen Union für Biochemie (IUB). Zusätzlich wurden die folgenden Abkürzungen benutzt.

% (v / v)	Volumenprozent
% (w / v)	Massenprozent
2-ODD	2-Oxoglutarat abhängige Dioxygenase
Amp	Ampicillin
APS	Ammoniumpersulfat
BCA	Bicinchoninsäure (<i>Bicinchoninic acid</i>)
BLOSUM	<i>Blocks Substitution Matrix</i>
bp	Basenpaare
BSA	Rinderalbumin (<i>Bovine Serum Albumin</i>)
c	Konzentration
cDNA	komplementäre DNA
Col-0	Columbia-0
CTAB	Cetylhexatrimethylammoniumbromid
CYP	Cytochrom-P450-abhängige Monooxygenase
d _{aa}	Aminosäuredistanz
dATP	Desoxyadenosintriphosphat
dCTP	Desoxycytidintriphosphat
DEAE	Diethylaminoethyl
dGTP	Desoxyguanosintriphosphat
DHS	Desoxyhypusinsynthase
DIG	Digoxigenin
DMF	Dimethylformamid
DMSO	Dimethylsulfoxid
d _N	Anzahl nichtsynonymer Substitutionen pro nichtsynonymer Sequenzposition
DNA	Desoxyribonukleinsäure
d _{na}	Nukleotiddistanz
dNTP	Desoxynukleosidtriphosphat
DOP	<i>Degenerate Oligonucleotide Primer</i>
d _s	Anzahl synonymer Substitutionen pro synonymer Sequenzposition
DTT	Dithiothreitol
dTTP	Desoxythymidintriphosphat
dUTP	Deoxyuridintriphosphat
EDTA	Ethylendiamintetraacetat
ESP	Epithiospezifizierendes Protein
GC-MS	Gaschromatographie-Massenspektrometrie
GFP	<i>Green Fluorescent Protein</i>
GH	Glycosidhydrolase
GSP	Genspezifischer Primer
GUS	Beta-Glucuronidase
HC	Hoch konzentriert (<i>Highly Concentrated</i>)

HPLC	Hochleistungsflüssigkeitschromatographie (<i>High-Pressure Liquid Chromatography</i>)
HRV 3C	3C-Protease des humanen Retrovirus Typ 14
HSS	Homospermidinsynthese
IPTG	Isopropyl- β -D-thiogalactopyranosid
Chl	Chloramphenicol
kb	Kilobasen
LB-Medium	Lysogeny-Broth-Medium (auch Luria-Bertani-Medium)
MAM	Methylthioalkylmalatsynthase
MES	2-Morpholinoethansulfonsäure-Monohydrat
Mill.	Millionen
M-MLV	Moloney Murine Leukemia Virus Reverse Transcriptase
n.t	Nicht getestet
n.i	Nicht identifizierbar
NBT	Nitroblau-Tetrazoliumchlorid
NSP	Nitril-spezififizierendes Protein
OD	Optische Dichte
ORF	Offener Leserahmen (<i>Open Reading Frame</i>)
PAM	<i>Percent Accepted Mutations-Matrix</i>
PCR	Polymerase-Kettenreaktion (<i>Polymerase Chain Reaction</i>)
Pfu-Polymerase	Polymerase aus <i>Pyrococcus furiosus</i>
RACE	<i>Rapid Amplification of cDNA-Ends</i>
RF	Response-Faktor
RNA	Ribonukleinsäure
RNAse	Ribonuclease
rpm	Umdrehungen pro Minute (<i>Revolutions Per Minute</i>)
RT	Reverse Transkriptase
SDS	Natriumdodecylsulfat
SOC-Medium	Super Optimal Broth-Medium
SOT	Sulfotransferase
SSC	Saline-Sodium-Citrate
Taq-Polymerase	Polymerase aus <i>Thermus aquaticus</i>
TBE-Puffer	Tris-Borat-EDTA-Puffer
TB-Medium	Terrific-Broth-Medium
TdT	Terminale Deoxynucleotidyltransferase
TEMED	N, N, N', N'-Tetramethylethylendiamin
TE-Puffer	Tris-EDTA-Puffer
TFP	Thiocyanat-formendes Protein
Tris	Tris(hydroxymethyl)-aminomethan
UGT	Thiohydroximat-S-Glucosyltransferase
USER	<i>Uracil-Specific Excision Reagent</i>
UTR	Untranslatierter Bereich (<i>Untranslated Region</i>)
UV	Ultraviolett
VF	Verdünnungsfaktor
X-Gal	5-Bromo-4-Chloro-3-Indolyl-beta-D-Galactopyranosid

1. EINLEITUNG

1.1 Das Glucosinolat-Myrosinase-System als Beispiel für ein pflanzliches Abwehrsystem.

Pflanzen produzieren eine immense Vielfalt von über 200.000 Sekundärmetaboliten, deren unverzichtbare Rolle als Komponenten der Überlebensstrategien von Pflanzen in den letzten Jahren große Anerkennung gefunden hat (Hartmann 2007). Während Strukturen des Primärstoffwechsels alle für das Wachstum und die Entwicklung der Pflanze essentiellen Funktionen abdecken, ist der Sekundärstoffwechsel wichtig für ein Überleben der Pflanzen unter den verschiedensten Umwelteinflüssen. Zu diesen zählen neben abiotischen auch biotische Faktoren wie Herbivorie und Parasitismus. Das auch als „Senfölbombe“ bezeichnete Glucosinolat-Myrosinase-System ist ein dem Sekundärstoffwechsel zuzuordnendes chemisches Abwehrsystem der Ordnung der Brassicales. Neben einer fraßhemmenden Wirkung auf Herbivoren beeinträchtigt es das Wachstum von Bakterien und Pilzen und besitzt eine Signalfunktion in der indirekten Abwehr (Wittstock et al. 2003; Kessler and Halitschke 2007; Mumm et al. 2008; Dicke and Baldwin 2010; Müller et al. 2010). Die Wirkung des Glucosinolat-Myrosinase-Systems beruht dabei auf der Hydrolyse der Glucosinolate, einer Gruppe anionischer und hydrophiler Thioglucoside. Die Glucosinolate kommen durch Verletzungen des pflanzlichen Gewebes, zum Beispiel durch Insektenfraß, in Kontakt mit ihren im intakten Gewebe räumlich getrennt vorliegenden hydrolytischen Enzymen, den Myrosinasen. Intakte Glucosinolate sind an sich ungiftig, sie werden jedoch nach Hydrolyse durch Myrosinasen spontan zu toxischen Abwehrstoffen, den Isothiocyanaten (Senfölen) abgebaut. In Anwesenheit spezifizierender Proteine können zudem Epithionitrile, organische Thiocyanate und einfache Nitrile als weitere biologisch aktive Verbindungen aus dieser Reaktion hervorgehen (Halkier and Gershenzon 2006). Im Laufe der Evolution konnten sich einige Herbivoren und Pathogene auf das Glucosinolat-Myrosinase-System einstellen (Winde and Wittstock 2011). Eine solche Anpassung führte zu Spezialisten wie dem Kohlweißling *Pieris rapae* (*P. rapae*; Lepidoptera: Pieridae), welcher die Glucosinolate sogar als Erkennungsmerkmal seiner Wirtspflanze nutzt (Renwick et al. 1992; Miles et al. 2005). Andere Herbivoren nutzen Pflanzen unterschiedlicher Familien mit divergierenden Sekundärstoffklassen. Dazu zählen Generalisten wie die Baumwolleneule *Spodoptera littoralis* (Lepidoptera: Noctuidae), welche auf einer Glucosinolat-haltigen Wirtspflanze überlebt, jedoch durch das Vorhandensein der Glucosinolate in ihrem Wachstum gehemmt wird (Müller et al. 2010). Besonders Spezialisten, aber auch Generalisten üben ihrerseits Selektionsdruck auf die Wirtspflanzen aus. Eine die Metaboliten des sekundären Stoffwechsels kennzeichnende Eigenschaft ist allerdings ihre hohe Formbarkeit und Vielfalt, welche durch die genetisch bedingte flexible Adaptation an die sich ständig ändernde Umwelt ermöglicht wird (Hartmann 2007) und so das Prinzip Coevolution (Ehrlich and Raven 1964) in eine neue Runde führt. Erste Anhaltspunkte weisen darauf hin, dass diese kontinuierliche Anpassung an wechselnde Umweltbedingungen unter anderem durch die Vervielfältigung und funktionelle Veränderung von Genen ermöglicht wird, die ursprünglich für Enzyme des primären Stoffwechsels codieren (Hartmann 2007). Ein gut untersuchtes Beispiel hierfür liefern die

Homospermidinsynthasen (HSS) aus der in entfernt verwandten Familien der Angiospermen vorkommenden Biosynthese der Pyrrolizidinalkaloide. HSS ist evolutiv durch Genduplikation aus einem Gen des Primärstoffwechsels entstanden, das für die Desoxyhypusinsynthase (DHS) codiert. Die DHS ist an der posttranslationalen Aktivierung eines regulatorischen Proteins, des eukaryotischen Translationsinitiationsfaktors 5A (eIF5A) beteiligt (Ober and Hartmann 1999a; Ober and Hartmann 1999b). HSS hat eine Nebenaktivität von DHS, nämlich die Fähigkeit Homospermidin aus Putrescin und Spermidin zu synthetisieren, als Hauptaktivität beibehalten, verlor hingegen die Affinität für die eIF5A-Vorstufe (Ober and Hartmann 2000; Ober et al. 2003). Auch innerhalb des Glucosinolat-Myrosinase-Systems konnten in den letzten Jahren erste abstammungsgeschichtliche Einsichten gewonnen werden, welche auf die Rekrutierung von Genen aus dem primären sowie aus dem sekundären Stoffwechsel hinweisen. So entstanden einzelne Proteine der Glucosinolat-Biosynthese aus Enzymen des Primärstoffwechsels (Benderoth et al. 2009; de Kraker and Gershenzon 2011), während andere aus Enzymen der dem sekundären Stoffwechsel zuzuordnenden Biosynthese der cyanogenen Glykoside abgeleitet sind (Kjaer 1973; Rodman et al. 1998; Mikkelsen et al. 2002; Bak et al. 2006). Neben der Biosynthese der Glucosinolate ist ihre Hydrolyse sehr bedeutsam für die Funktion des Glucosinolat-Myrosinase-Systems. Bei der Analyse der hydrolytischen Enzyme ergeben sich ebenfalls Parallelen zu den cyanogenen Glykosiden (Rask et al. 2000; Xu et al. 2004). Zusätzlich haben spezifizierende Proteine einen großen Einfluss auf die Vielfalt der aus den Glucosinolaten entstehenden Glucosinolat-Hydrolyseprodukten, da in ihrer Anwesenheit neben den spontan gebildeten Isothiocyanaten ebenfalls einfache Nitrile, Epithionitrile und organische Thiocyanate entstehenden. Auf Grund der nur in geringer Anzahl bekannten spezifizierenden Proteine konnten bisher jedoch nur erste spekulative Aussagen über ihre evolutive Abstammung gemacht werden, welche noch viele Fragen offen lassen.

1.2 Glucosinolat-Biosynthese und ihr evolutiver Hintergrund

Die Grundstruktur der Glucosinolate (Abbildung 1 A) wurde erstmals 1976 von Kjær und Olsen Larsen vorgeschlagen (Kjær and Olsen Larsen 1976). Über 120 Strukturen konnten innerhalb der Ordnung der Brassicales nachgewiesen werden. Sie gehen in ihrer Biosynthese aus nur acht proteinogenen Aminosäuren hervor. Entsprechend ihrer Herkunft können die Glucosinolate in die drei Gruppen aliphatisch (aus Methionin, Alanin, Valin, Leucin, Isoleucin), aromatisch (aus Tyrosin, Phenylalanin) und indolisch (aus Tryptophan) unterteilt werden (Fahey et al. 2001) (Abbildung 1).

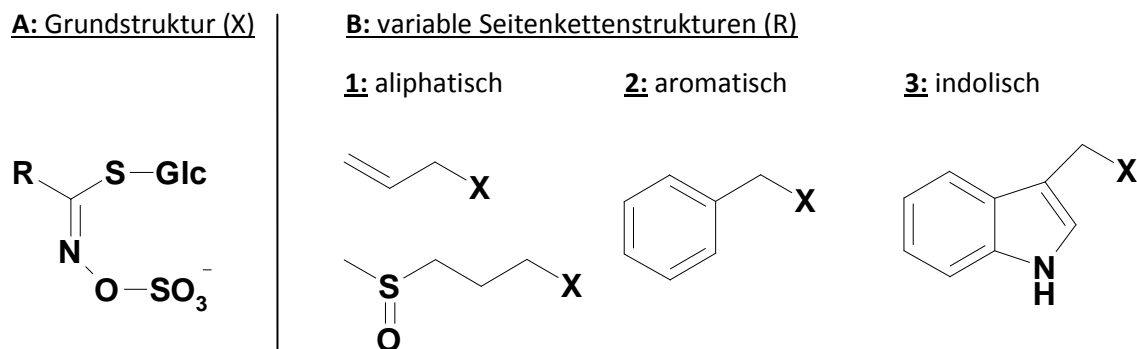


Abbildung 1: Chemische Struktur der Glucosinolate. Die Grundstruktur der Glucosinolate (A) besteht aus einem (Z)-N-Hydroximinosulfatester der mit einer thioglucosidisch gebundenen Glucose und einer variablen Seitenkette verknüpft ist. Die variablen Seitenkettenstrukturen (B) gliedern die Glucosinolate je nach ihrer Ausgangsaminosäure in aliphatische (1) (z.B. Allylglucosinolat (oben) und 4-Methylsulfinylbutylglucosinolat (unten)), aromatische (2) (z.B. Benzylglucosinolat) und indolische (3) Glucosinolate (z.B. Indol-3-ylmethylglucosinolat).

Modifiziert nach: Winde and Wittstock 2011

Die Glucosinolat-Biosynthese unterteilt sich in drei Abschnitte (Abbildung 2): (I) Kettenverlängerung der Ausgangsaminosäure (nur bei Methionin und Phenylalanin), (II) Bildung der Glucosinolatgrundstruktur (Abbildung 1) und (III) sekundäre Modifikation der Glucosinolatseitenkette. Das Glucosinolatprofil einzelner Spezies ist jeweils spezifisch (Fahey et al. 2001), kann jedoch selbst innerhalb der Individuen einer Art in Gehalt und Zusammensetzung, bedingt durch biotische und abiotische Faktoren, sowie im Verlauf der pflanzlichen Entwicklung und innerhalb einzelner Pflanzengewebe sowie -organe variieren (Brown et al. 2003; Charron et al. 2005; Charron et al. 2005; Himanen et al. 2008; Wentzell and Kliebenstein 2008; Textor and Gershenzon 2009). Innerhalb der letzten Jahre konnten in *Arabidopsis thaliana* (*A. thaliana*; Brassicaceae) mehr als 20 Gene identifiziert werden, welche maßgeblich für die Biosynthese der Glucosinolate verantwortlich sind. Ihre Identifizierung eröffnete die Möglichkeit zur Untersuchung ihrer evolutiven Abstammung. Am besten untersucht ist die Biosynthese aliphatischer Glucosinolate aus *A. thaliana*.

1.2.1 Schritt I: Kettenverlängerung der Ausgangsaminosäure

Eingeleitet wird die in mehreren Zyklen ablaufende Kettenverlängerung der Aminosäure Methionin in *A. thaliana* durch eine im Cytosol stattfindende Desaminierung des Methionins durch die Aminotransferase BCAT4 (Abbildung 2 ①). Das entstehende Produkt wird über den Transporter BAT5 (Gigolashvili et al. 2009) in den Chloroplasten eingeschleust und unterliegt dort einer Kondensation mit Acetyl-CoA (Abbildung 2 ①). Die an dieser Kondensation beteiligten Methylthioalkylmalatsynthasen MAM1 (im ersten, zweiten und dritten Zyklus) und MAM2 (nur im ersten Zyklus) scheinen durch eine Genduplikation von *MAMa* entstanden zu sein. *MAMa*, welches sich unter anderem in *Arabidopsis lyrata* (*A. lyrata*; Brassicaceae) und *Boechea divaricarpa* (Brassicaceae) identifizieren ließ, ist Teil eines Clusters aus drei *MAM*-Genen, *MAMa*, *MAMb* und *MAMc*, welche ebenfalls aus einer > 10 Millionen Jahre (Mill. Jahre) zurückliegenden Duplikation entstanden sind (Benderoth et al. 2006). Phylogenetische Analysen aller *MAM*-Gene konnten zeigen, dass die Duplikation von *MAMa* nach der Entstehung von *A. thaliana* stattgefunden hat (vor ~13 Mill. Jahren (Abbildung 5) (Benderoth et al. 2009; Beilstein et al. 2010). Die dritte bekannte Methylthioalkylmalatsynthase MAML, heute als MAM3 bezeichnet, unterstützt an diesem Punkt der Kettenverlängerung des Methionins die Bildung längerer Seitenketten (Abbildung 2 ①) (Benderoth et al. 2006) und konnte als Ortholog von MAMb identifiziert werden (Benderoth et al. 2009). Allen *MAM*-Genen aus *A. thaliana* gemeinsam ist ihre Abstammung aus Isopropylmalatsynthasen (IPMS) des primären Stoffwechsels mit katalytischer Funktion in der Leucin-Biosynthese. Der Ursprung der *MAM*-Gene aus den *IPMS*-Genen zeigt sich an Hand hoher Sequenzähnlichkeiten, phylogenetisch naher Verwandtschaft und einer großen strukturellen Ähnlichkeit (Benderoth et al. 2009; de Kraker and Gershenzon 2011). Die weiteren Schritte der Kettenverlängerung im Fall von Methionin verlaufen über eine Isomerisierung durch eine Isopropylmalat-Isomerase (IPMI), eine oxidative Decarboxylierung durch eine Isopropylmalat-Dehydrogenase (IPMDH1), gefolgt von einer Transaminierung vermutlich durch die Aminotransferase BCAT3 und einem anschließenden Eintritt in die Bildung der Glucosinolatgrundstruktur die am Endoplasmatischen Retikulum im Cytosol stattfindet (Abbildung 2 ①). Alternativ kann das Molekül im Chloroplasten verbleiben und bis zu acht weitere Zyklen der Kettenverlängerung durchlaufen (Abbildung 2 ①) (Halkier and Gershenzon 2006).

1.2.2 Schritt II: Bildung der Glucosinolatgrundstruktur

Weitere Hinweise auf die evolutive Abstammung einzelner Proteine der Glucosinolat-Biosynthese finden sich in der Bildung der Glucosinolatgrundstruktur. In *A. thaliana* gehen in diesen Abschnitt kettenverlängerte Methionin-Derivate (Homomethionin bis Hexahomomethionin), sowie Phenylalanin, Homophenylalanin und Tryptophan ein (Abbildung 2 ②). In einem ersten Schritt werden diese Ausgangsstoffe durch Cytochrom-P450-abhängige Monooxygenasen (CYP) der CYP79-Familie in Aldoxime umgewandelt (Abbildung 2 ②). Dabei katalysiert CYP79A2 die Umwandlung des Phenylalanins, CYP79B2 und CYP79B3 die des Tryptophans und CYP79F1 und CYP79F2

die der Methionin-Derivate (Abbildung 2 ②). CYP79B2 und CYP79B3 unterscheiden sich in der Genregulation und CYP79F1 und CYP79F2 in ihrer Spezifität für Methionin-Derivate unterschiedlicher Kettenlängen. Die Funktion der nah verwandten CYP79C1 und CYP79C2 konnte bisher nicht zugeordnet werden (Abbildung 2 ②). Es wird jedoch angenommen, dass sie auf Grund ihres geringen Transkriptionslevels an der Synthese von nur in geringen Mengen vorkommenden Glucosinolaten beteiligt sind (Mikkelsen et al. 2002). Schon 1973 wurde erstmals die Hypothese aufgestellt, dass Enzyme der Glucosinolat-Biosynthese aus Vorläufern entstanden sind, die an der Biosynthese der cyanogenen Glycoside beteiligt sind (Kjaer 1973). Diese Annahme erhärtete sich durch die Isolierung von CYPs z.B. aus *A. thaliana*, *Brassica napus* (*B. napus*; Brassicaceae), *Carica papaya* (Caricaceae), *Sinapis alba* (*S. alba*; Brassicaceae) und *Tropaeolum majus* (Tropaeolaceae) (Bak et al. 1998) sowie durch die Entdeckung einer nahen phylogenetischen Verwandtschaft von Pflanzen mit cyanogenen Glycosiden, wie beispielsweise Spezies der Familien Euphorbiaceae, Passifloraceae, und Sapindaceae zu Spezies, welche sich auf Grundlage des Glucosinolat-Myrosinase-Systems verteidigen (Rodman et al. 1998).

Die in diesem Biosyntheseschritt entstandenen Aldoxime werden im weiteren Verlauf durch Enzyme der CYP83-Familie zu reaktiven Verbindungen oxidiert (Abbildung 2 ②). Hierbei wird die Oxidation der von Tryptophan und Phenylalanin abgeleiteten Aldoxime durch CYP83B1 und die aller aliphatischen Aldoxime von CYP83A1 katalysiert (Abbildung 2 ②). Auf Aminosäureebene haben CYP83A1 und CYP83B1 eine Identität von 55 %. Des Weiteren zeigen sie eine große Sequenzähnlichkeit zu CYP71E1 aus *Sorghum bicolor* (Poaceae), einem CYP, das an der Biosynthese des cyanogenen Glycosids Dhurrin beteiligt ist. Auf Grund ihrer Sequenzähnlichkeit gruppieren sich CYP71E1 sowie weitere Cytochrome der CYP71-Familie in phylogenetischen Analysen gemeinsam mit Mitgliedern der CYP83-Familie, was deutlich die Zusammengehörigkeit beider Familien aufzeigt. Ungeklärt ist jedoch, ob die Vertreter der CYP83-Familie der Glucosinolat-Biosynthese aus einem CYP71E1 Ortholog hervorgegangen sind, oder ob andere CYP71 Homologe einbezogen wurden (Bak et al. 2006). Ebenfalls im Dunkeln liegt der Zeitpunkt dieser evolutiven Ereignisse. Eine Abstammung der CYPs der Glucosinolat-Biosynthese von denjenigen der Synthese cyanogener Glycoside scheint jedoch klar. Diskutiert wird eine Mutation in einem ursprünglichen CYP-Gen, die zur Bildung eines für die Pflanze giftigen oxidierten Aldoxims führte, statt zu den Aldoxim-Intermediaten der Biosynthese der cyanogenen Glycoside. Das oxidierte Aldoxim musste entgiftet werden. Dies könnte dann durch Konjugation mit einem Schwefeldonor, Glycosylierung und Sulfatierung zu den heute bekannten Glucosinolaten geführt haben (Hansen et al. 2001). Anzumerken ist, dass die Fähigkeit cyanogene Glycoside zu erzeugen mindestens 300 Mill. Jahre alt zu sein scheint, da diese Verbindungen nicht nur in Angiospermen (140 Mill. Jahre alt) und Farnen (320 Mill. Jahre alt) sondern auch in Gymnospermen (360 Mill. Jahre alt) identifiziert werden konnten (Bak et al. 2006).

Der oben aufgezeigte mögliche Entgiftungsweg der oxidierten Aldoxime konnte in den letzten Jahren durch experimentelle Arbeiten als Biosyntheseweg der Glucosinolaten bestätigt werden (Sonderby et al. 2010). Die aktivierten Aldoxime werden im Folgenden mit einem Schwefeldonor konjugiert (Abbildung 2 ②). Lange Zeit wurde Cystein als wahrscheinlichster

Schwefeldonor diskutiert. Neueste Studien haben jedoch Glutathion als Schwefeldonor in der Glucosinolat-Biosynthese identifiziert (Geu-Flores et al. 2011). Diese Tatsache macht einen weiteren Schritt, katalysiert durch Glutamyl-Transferasen (GPP1 und GPP3), notwendig (Abbildung 2 ②). Die entstandenen Glutathionkonjugate unterliegen einer Spaltung durch die C-S-Lyase SUR1 (Abbildung 2 ②). Es entstehen Thiohydroximsäuren, welche wiederum durch Thiohydroximat-S-Glucosyltransferasen (UGT) der UGT74-Familie S-glucosyliert werden, was zur Bildung der Desulfoglucosinolate führt (Abbildung 2 ②). Bei diesem Schritt katalysiert UGT74B1 vermutlich die Glucosylierung der aus Phenylalanin abgeleiteten Thiohydroximate, UGT74C1 hingegen die der aus Methionin abgeleiteten Thiohydroximate (Abbildung 2 ②) (Sonderby et al. 2010). Auch hier bestehen hohe Sequenzidentitäten dieser UGTs zu Glucosyltransferasen aus der Biosynthese cyanogener Glycoside, was eventuell auf eine Entstehung durch eine Genduplikation hinweisen könnte (Mikkelsen et al. 2002). Die aus der Glucosylierung hervorgegangenen Desulfoglucosinolate werden anschließend durch die Sulfotransferasen (SOT) SOT16, SOT17 und SOT18 zu Glucosinolaten sulfatiert (Abbildung 2 ②). Die aus Tryptophan und Phenylalanin abgeleiteten Desulfoglucosinolate sind Substrate von SOT16, die aus Methionin abgeleiteten Desulfoglucosinolate sind Substrate von SOT17 und SOT18 (Abbildung 2 ②).

1.2.3 Schritt III: Sekundäre Modifikation der Glucosinolatseitenkette

Die dritte Phase der Glucosinolat-Biosynthese ist die nachträgliche (sekundär genannte) Modifikation der Glucosinolatseitenkette, welche einen großen Beitrag zur Bildung der immensen Vielfalt der Glucosinolate leistet. Eine Möglichkeit in *A. thaliana* besteht hier in der Anreicherung aliphatischer Glucosinolate mit Sauerstoff (Oxygenierung) durch die Flavin-abhängige Monooxygenasen FMO_{GS-OX1} bis FMO_{GS-OX5}, welche sowohl kurz- als auch langkettige aliphatische Glucosinolate oxygenieren können (Abbildung 2 ③) (Sonderby et al. 2010). Hierbei ist FMO_{GS-OX5} spezifisch für die Reaktion mit langkettigen, aliphatischen Glucosinolaten, während FMO_{GS-OX1} bis FMO_{GS-OX4} Glucosinolate mit unterschiedlich langen Seitenketten akzeptieren (Abbildung 2 ③) (Hansen et al. 2007; Li et al. 2008). Methylsulfinylalkylglucosinolate entstehen in *A. thaliana* durch die Konvertierung der aliphatischen und mit Sauerstoff angereicherten Glucosinolate zu Alkenylglucosinolaten durch die 2-Oxoglutarat abhängige Dioxygenase (2-ODD) AOP2 (Abbildung 2 ③). AOP3 wandelt 3-Methylsulfinylpropylglucosinolat in 3-Hydroxypropylglucosinolat um (Abbildung 2 ③). Ein drittes AOP-Gen (AOP1) in *Arabidopsis* konnte bisher in seiner Funktion nicht charakterisiert werden. Eine Genduplikation von AOP1 scheint jedoch zu einem Vorläufer von AOP2 und AOP3 geführt zu haben. Eine weitere Genduplikation dieser Vorläufersequenz ermöglichte dann die Differenzierung zu diesen zwei Varianten der 2-ODDs. Bei der zweiten Genduplikation wird vermutet, dass sie lange vor der Aufspaltung von *A. thaliana* und *A. lyrata* erfolgte (vor > 13 Mill. Jahren (Abbildung 5)) (Kliebenstein et al. 2001).

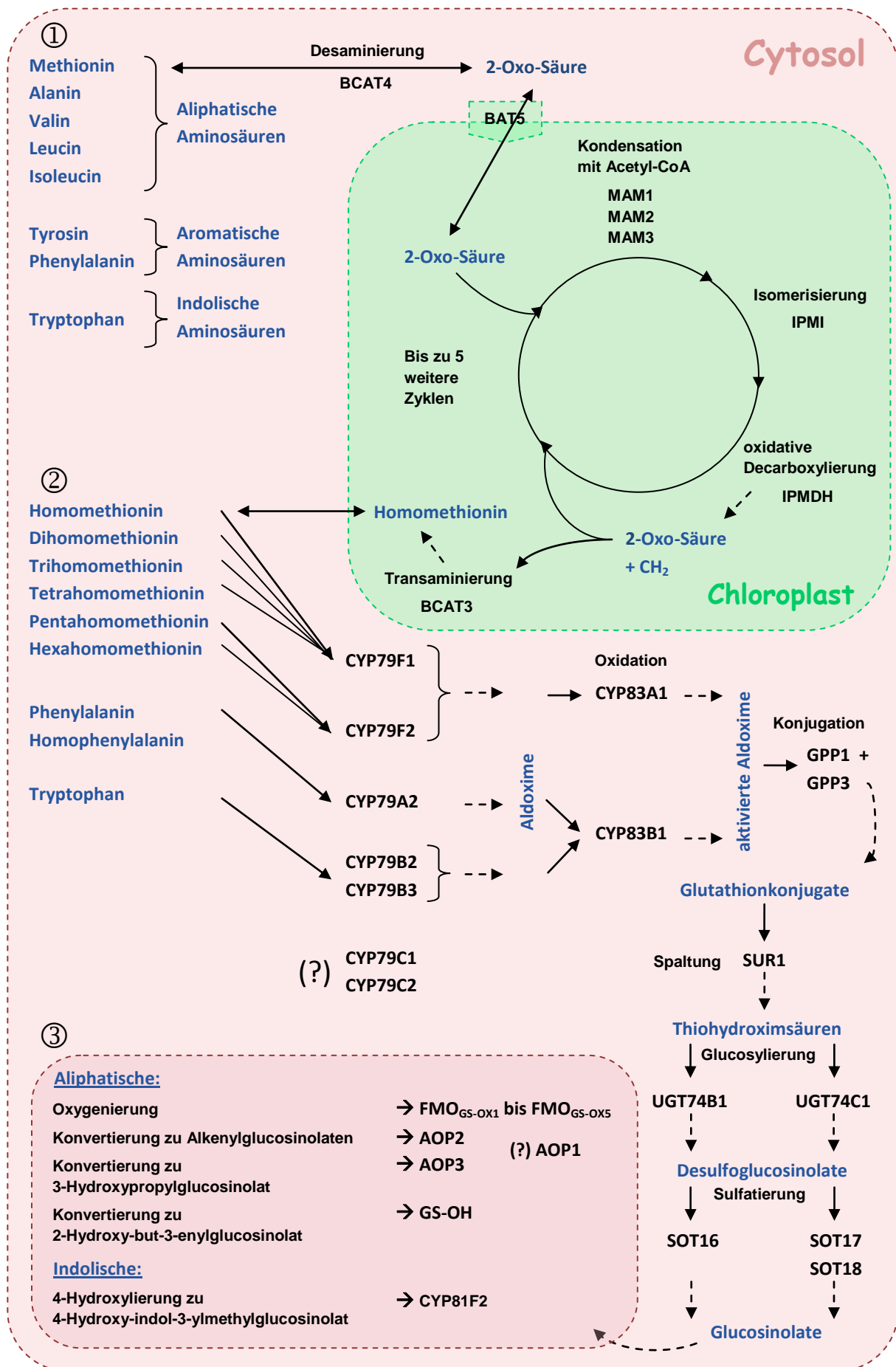
Eine andere Gruppe von aliphatischen Glucosinolaten in *A. thaliana*, die Hydroxyalkenylglucosinate, werden ebenfalls durch 2-ODDs gebildet (Abbildung 2 ③). Experimentell nachgewiesen ist die Umsetzung von 3-Butenylglucosinolat zu 2-Hydroxy-but-3-enylglucosinolat durch die 2-ODD GS-OH (Abbildung 2 ③) (Hansen et al. 2008). Eine weitere sekundäre Modifikation der Glucosinolatseitenkette ist die Benzoylierung hydroxylierter aliphatischer Glucosinate. Zugehörige Proteine konnten jedoch bisher nicht identifiziert werden. Über die sekundären Modifikationen der Glucosinolatseitenkette indolischer Glucosinate ist weit weniger bekannt. Hier konnte in *Arabidopsis* bisher als einziges Enzym ein CYP der CYP81-Familie, das CYP81F2 für die 4-Hydroxylierung von Indol-3-ylmethylglucosinolat zu 4-Hydroxy-indol-3-ylmethylglucosinolat verantwortlich gemacht werden (Abbildung 2 ③) (Pfalz et al. 2009).

1.3 Kompartimentierung

In pflanzlichen Geweben liegen Glucosinate und Myrosinasen in getrennten Kompartimenten vor. Glucosinate kommen im gesamten pflanzlichen Gewebe vor, in besonders hohen Konzentrationen aber vermutlich in den sogenannten S-Zellen, die dadurch einen hohen Schwefelgehalt aufweisen (Koroleva et al. 2000). S-Zellen säumen das Phloem zwischen den Leitbündeln und der Epidermis. Myrosinasen konnten in *Arabidopsis* in Idioblasten des Phloem-Parenchyms, die auch Myrosinzellen genannten werden, (Xue et al. 1995; Andreasson et al. 2001) sowie in Schließzellen und Phloem-assoziierten Zellen oberirdischer Pflanzenteile nachgewiesen werden (Husebye et al. 2002; Thangstad et al. 2004; Barth and Jander 2006). In Keimlingen und Samen von *Brassica juncea* (Brassicaceae) kommen Glucosinate und Myrosinasen gemeinsam innerhalb aleuronähnlicher Zellen vor. In diesem Fall wird von einer Kompartimentierung auf subzellulärer Ebene ausgegangen (Kelly et al. 1998).

Kommen Glucosinate und Myrosinasen durch Gewebeverletzungen in Kontakt, so wird die Thioglucosidbindung des Glucosinolates gespalten und Glucose freigesetzt. Dabei entsteht ein instabiles Aglycon (Abbildung 3). Durch Lossen-ähnliche-Umlagerung werden aus dem Aglycon spontan die toxischen Isothiocyanate gebildet.

Abbildung 2: Glucosinolat-Biosynthese. Dargestellt ist die Biosynthese aliphatischer und indolischer Glucosinate, gegliedert in die Kettenverlängerung der Ausgangsaminosäure (1), die Bildung der Glucosinolatgrundstruktur (2) und die sekundäre Modifikation der Glucosinolatseitenkette (3). Abgebildet wurden nur sicher identifizierte sowie partiell charakterisierte Enzyme. Nähere Erläuterungen zu Metaboliten und enzymatischen Schritten sind im zugehörigen Textabschnitt ausgeführt. Modifiziert nach: Sonderby et al. 2010



1.4 Myrosinasen und ihr evolutiver Hintergrund

Die hydrolytischen Enzyme des Glucosinolat-Myrosinase-Systems sind die als Myrosinasen bezeichneten Thioglucosidasen. Die erste Myrosinase wurde bereits 1840 von A. Bussy in *Brassica nigra* (Brassicaceae) beschrieben (Bussy 1840), aber auch in weiteren Glucosinolat-haltigen Spezies der Familie Brassicaceae konnten im Lauf der Jahre Myrosinasen identifiziert werden (z.B. *Lepidium sativum* (Durham and Poulton 1990), *S. alba* und *B. napus* (Bjorkman and Lonnerdal 1973)). Die vier in *A. thaliana* detektierten Myrosinasen TGG1 (At5g26000, BGLU38), TGG2 (At5g25980, BGLU37), TGG4 (At1g47600, BGLU34) und TGG5 (At1g51470, BGLU35) (Xu et al. 2004) zählen zu den klassischen Myrosinasen und konnten zusammen mit weiteren prokaryotischen und eukaryotischen Vertretern der Glycosidhydrolase (GH) Familie 1 zugeordnet werden (Rask et al. 2000). Außerdem konnten zwei weitere für Myrosinasen codierenden Gene, TGG3 (At5g48375, BGLU39) und TGG6 (At1g51490, BGLU36), als Pseudogene in die GH-Familie 1 eingeordnet werden. Die beiden Pseudogene werden exprimiert, besitzen jedoch keine Funktionalität (Zhang et al. 2002; Xu et al. 2004; Wang et al. 2009). Innerhalb der GH-Familie 1 bilden die Myrosinasen ein zusammenhängendes phylogenetisches Cluster, was vermuten lässt, dass sich die Myrosinasen evolutiv von O-Glycosidasen ableiten und eine Spezialisierung auf die Hydrolyse der Glucosinolate durchlaufen haben (Xu et al. 2004). Auch die cyanogenen O- β -Glycosidasen gehören zur GH-Familie 1. Eindeutige Aussagen über eine Abstammung der Myrosinasen aus dem System der cyanogenen Glycoside basierend auf phylogenetischen Analysen charakterisierter Myrosinasen und pflanzlicher Vertreter cyanogener und nicht-cyanogener O- β -Glycosidasen der GH-Familie 1 konnten jedoch nicht getroffen werden (Rask et al. 2000).

Strukturelle Ähnlichkeiten innerhalb der Myrosinasen und cyanogenen β -Glycosidasen geben jedoch Hinweise auf eine Ableitung der Myrosinasen aus cyanogenen β -Glycosidasen (Rask et al. 2000). Im Vergleich der Linamarase, einer cyanogenen β -Glycosidase aus *Triticum repens* (Poaceae), und der Myrosinase aus *S. alba* zeigen beide Glycosidasen eine Organisation als Homodimer, ein breites Substratspektrum, eine identische Größe, eine ähnliche dreidimensionale Struktur und einen ähnlichen Reaktionsmechanismus (Rask et al. 2000). Des Weiteren weisen die Myrosinasen TGG1 bis TGG3 eine Genstruktur aus 12 Exons und 11 Introns auf. Ein Vergleich der Exon / Intron-Struktur von TGG1 als beispielhaftem Vertreter der Myrosinasen und drei weiterer charakterisierter pflanzlicher O- β -Glycosidasen unterschiedlicher Spezifitäten aus *Manihot esculenta* (bglA, X94986 (Liddle et al. 1998)), *Zea mays* (glu1, U44773 (Brzobohaty et al. 1993)) und *Hordeum vulgare* (bgq60, L41869 (Leah et al. 1995)) zeigte, dass die Anordnung der Genstruktur in TGG1 bis auf wenige Fusionen und Aufspaltungen mit der Genstruktur von β -O-Glycosidasen übereinstimmt (Rask et al. 2000). Da eine derart ähnliche Entwicklung unabhängig voneinander extrem unwahrscheinlich ist bestätigt die nachgewiesene Exon / Intron-Struktur ebenfalls die Annahme eines gemeinsamen evolutiven Hintergrundes von Myrosinasen und O- β -Glycosidasen (Rask et al. 2000).

1.5 Spezifizierende Proteine als Teil des Glucosinolat-Myrosinase-Systems

Anfänglich wurde das Glucosinolat-Myrosinase-System als Zwei-Komponenten-System, bestehend aus den an sich ungiftigen Glucosinolaten und den Myrosinasen beschrieben (Matile 1980). Jedoch sind seit ca. 50 Jahren Hinweise auf die Beteiligung von spezifizierenden Proteinen bekannt (Gmelin 1959; Tookey 1973; Benn 1977) und konnten kürzlich durch die Klonierung und Charakterisierung des ersten Thiocyanat-formenden Proteins (TFP) aus *Lepidium sativum* (Brassicaceae) (LsTFP) sowie die Identifizierung der Epithiospezifizierenden Proteine (ESP) aus *A. thaliana* Ökotyp *Landsberg erecta* (AtESP) und *Brassica oleracea* ssp. *oleifera* var. *italica* (Brassicaceae) (BoESP) sowie der fünf Nitril-spezifizierenden Proteine (NSP) aus *A. thaliana* Ökotyp Columbia-0 (AtNSP1 - AtNSP5) bestätigt werden (Tabelle 1). Desweiteren konnte der Familie der spezifizierenden Proteine in *A. thaliana* eine ähnliche Sequenz zugeordnet werden, für die jedoch keine Aktivität als spezifizierendes Protein der Glucosinolat-Myrosinase-Reaktion nachgewiesen werden konnte und für die aus diesem Grund nur der aus dem Genlokus abgeleitete Name At3g07720 (NM_556316) angegeben wird (Burow et al. 2009; Kissen and Bones 2009).

Tabelle 1: Biochemisch charakterisierte spezifizierende Proteine. Aufgeführt sind spezifizierende Proteine, die im Vorfeld bekannt waren. Name, Genlokus, Accession-Nummer und Referenzen sind angegeben.

Name	Genlokus	Accession-nummer	Referenz
LsTFP	---	DQ417116	(Burow et al. 2007a)
BoESP	---	DQ059298	(Matusheski et al. 2006)
AtESP	At1g54040	NP_175806	(Lambrix et al. 2001)
AtNSP1	At3g16400	NP_566546	(Burow et al. 2009; Kissen and Bones 2009)
AtNSP2	At2g33070	NP_180866	(Burow et al. 2009; Kissen and Bones 2009)
AtNSP3	At3g16390	NP_566545	(Burow et al. 2009; Kissen and Bones 2009)
AtNSP4	At3g16410	NP_188262	(Burow et al. 2009; Kissen and Bones 2009)
AtNSP5	At5g48180	NP_568692	(Burow et al. 2009; Kissen and Bones 2009)

Spezifizierende Proteine sind neben der Variabilität der Glucosinolat-Biosynthese für die Vielfalt der entstehenden Glucosinolat-Hydrolyseprodukte verantwortlich. Abhängig von den Reaktionsbedingungen, strukturellen Voraussetzungen der Glucosinolate und der Anwesenheit spezifizierender Proteine können aus der Glucosinolat-Myrosinase-Reaktion neben Isothiocyanaten auch Epithionitrile, organische Thiocyanate oder einfache Nitrile hervorgehen (Abbildung 3). Die spezifizierenden Proteine haben dabei keine hydrolytische Aktivität in Bezug auf die Glucosinolate. Die spezifizierenden Proteine agieren vermutlich als Enzyme, welche das Aglycon zu den alternativen Produkten umsetzen (Wittstock and Burow 2007). Hierbei fordert die Bildung von Epithionitrilen neben der Anwesenheit eines ESP eine

endständige Doppelbindung in der Glucosinolatseitenkette sowie das Vorhandensein von Eisenionen (Burow et al. 2006a). Aus Glucosinolaten, die keine Doppelbindung aufweisen entstehen in Anwesenheit eines ESP einfache Nitrile (Abbildung 3). Die Bildung von organischen Thiocyanaten ist nur in Gegenwart eines TFP möglich, welches außerdem die Fähigkeit zur Bildung von Epithionitrilen und einfachen Nitrilen besitzt (Burow et al. 2007a; Kuchernig et al. 2011) (Abbildung 3). NSPs erlauben alleinig die Umsetzung zu einfachen Nitrilen (Burow et al. 2009) (Abbildung 3). Einfache Nitrile bilden zudem einen Spezialfall, da sie auch in Abwesenheit spezifizierender Proteine bei pH-Werten unter fünf oder bei hohen Eisenionenkonzentrationen entstehen (Ettlinger et al. 1961; Tookey and Wolff 1970; Burow et al. 2006a).

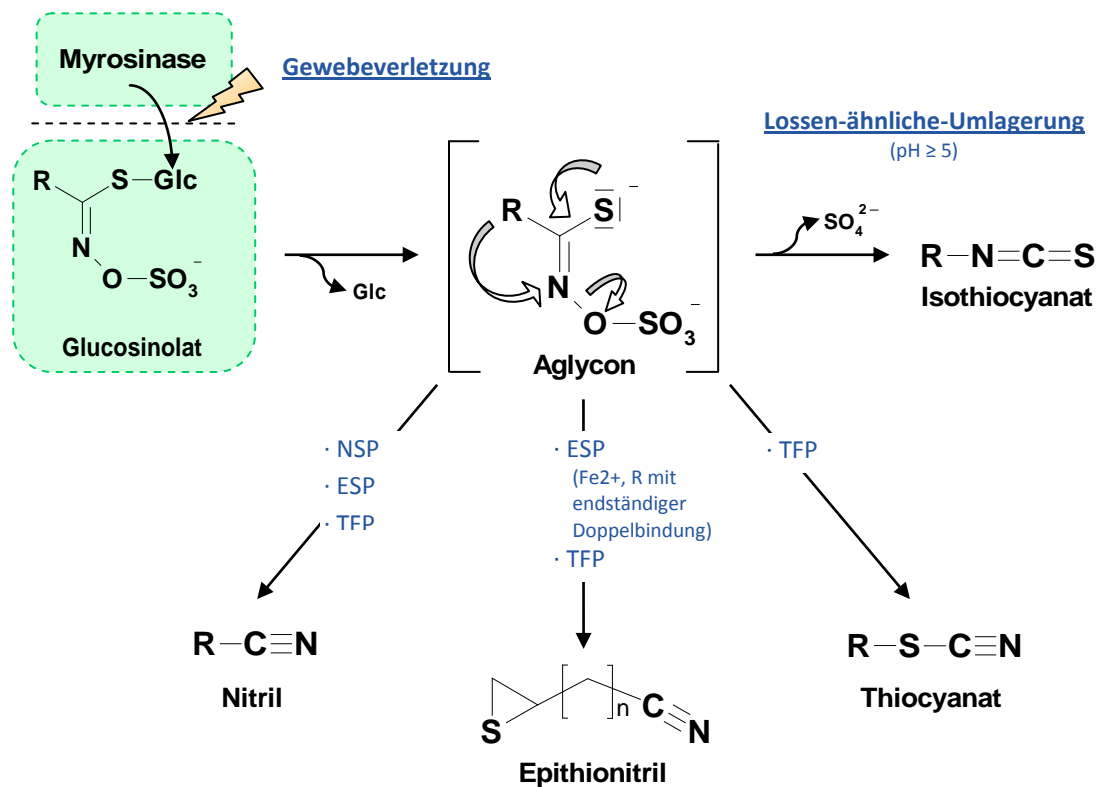


Abbildung 3: Glucosinolat-Myrosinase-System. Durch Gewebeverletzung, z.B. durch Insektenfraß, kommen Glucosinolatside und Myrosinasen in Kontakt. Dies führt zu einer hydrolytischen Abspaltung der Glucose und der Freisetzung eines instabilen Aglycons, welches sich spontan zu Isothiocyanaten umlagert. Die Bildung der alternativen Hydrolyseprodukte (einfache Nitrile, Epithionitrile und organischen Thiocyanate) setzt die Anwesenheit der angegebenen spezifizierenden Proteine voraus. Notwendige ergänzende Faktoren stehen in Klammern unter dem entsprechenden spezifizierenden Protein.

Strukturelle Betrachtungen der bisher nur in geringer Anzahl bekannten Sequenzen charakterisierter spezifizierender Proteine (Tabelle 1) und At3g07720 ergaben in aktuellen Analysen das Vorhandensein einer Abfolge von drei (AtNSP5 und At3g07720) bis vier (AtESP, BoESP, LsTFP und AtNSP1 – AtNSP4) Kelch-Motiven. Jedes Kelch-Motiv repräsentiert ein vier-strängiges β -Faltblatt das ein einzelnes Blatt eines Propellers darstellt. Die Wiederholung von Kelch-Motiven (Kelch-Wiederholungen) in einer Sequenz ergibt somit eine konservierte Proteinstruktur die β -Propeller genannt wird. Das Kelchmotiv wurde nach einer *Drosophila*-Mutante (Diptera: Drosophilidae) benannt, in welcher es erstmals identifiziert werden konnte (Xue and Cooley 1993). Die Beteiligung von Kelch-Proteinen an einer Vielzahl von Protein-Protein-Interaktionen sowohl innerhalb als auch außerhalb der Zelle gilt als experimentell bestätigt. Sie beinhaltet unter anderem die Assoziation mit Actin, die Regulation der Genexpression und eine extrazelluläre Rolle (Adams et al. 2000). Im Fall von AtNSP1, AtNSP2 und AtNSP3 ist den Kelchmotiven N-terminal jeweils eine Lektin-ähnlichen Jacalindomäne vorangestellt (Abbildung 4). AtNSP4 wird von zwei aufeinander folgenden N-terminalen Lektin-ähnlichen Jacalindomänen angeführt (Abbildung 4). AtNSP5 sowie At3g07720 besitzen ebenso wie die bisher bekannten *ESP*- und *TFP*-Sequenzen keine Lektin-ähnliche Jacalindomäne (Burow et al. 2009) (Abbildung 4). Lektine sind Proteine, die die Fähigkeit besitzen, Zucker zu binden sowie Glykokonjugate zu präzipitieren und Zellen zu agglutinieren. Jacalin ist ein Lektin, bekannt aus der Jackfrucht (*Artocarpus integrifolia*), das ein für Galaktose spezifisches Agglutinin darstellt (Jeyaprakash et al. 2002).

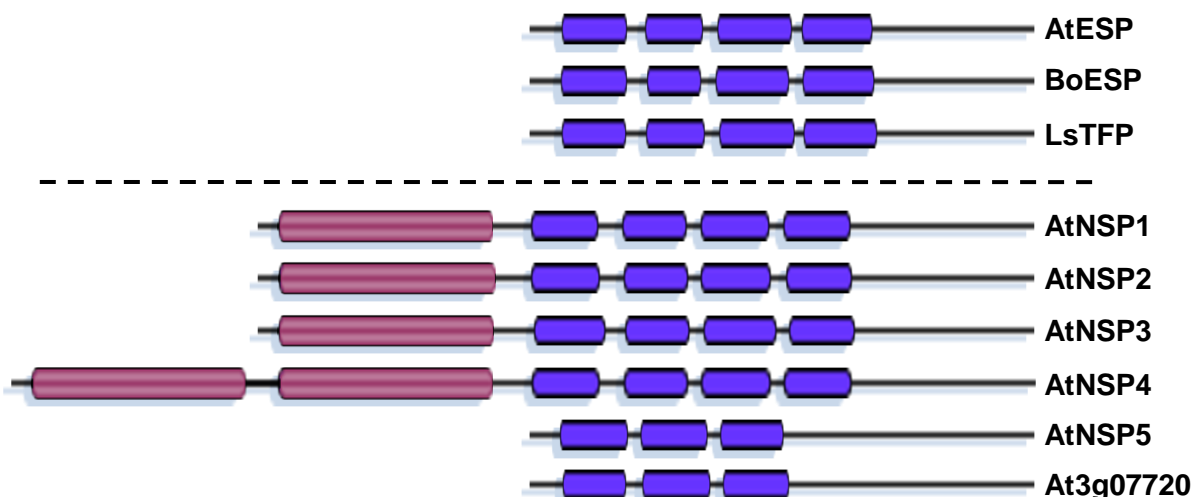


Abbildung 4: Vorhersage der Domänenstruktur bekannter spezifizierender Proteine. Mittels InterProScan (EMBL-EBI Webserver) identifizierte Kelchdomänen (blaue Kästen) und Lektin-ähnliche Jacalindomänen (violette Kästen) der bisher bekannten spezifizierenden Proteine (Tabelle 1) sowie von At3g07720 wurden dargestellt.

Die 2009 von Burow et al. erstellten Vergleiche der bekannten spezifizierenden Proteine (Tabelle 1) zeigten eine Aminosäuresequenzidentität von 50 % bis 60 % zwischen AtNSPs und AtESP. Die Identität der Aminosäuresequenzen von AtESP, BoESP und LsTFP liegt mit 63 % - 67 % höher (Burow et al. 2009). Die betrachteten Lektin-ähnlichen Jacalindomänen der AtNSPs weisen eine Ähnlichkeit von 50 % bis 52 % auf Aminosäureebene zu Sequenzen aus *A. thaliana* auf, die für die sogenannten Myrosinase-bindenden Proteine (MBP) (MBP1 (At1g52040, NP_175616) und MBP2 (At1g52030, NP_849794) (Theologis et al. 2000)) codieren (Burow et al. 2009). MBPs bestehen fast ausschließlich aus Jacalindomänen (Falk et al. 1995; Taipalensuu et al. 1997; Nagano et al. 2008) und sind eine Gruppe von Proteinen, welche in Spezies der Gattung Brassica als Komplexbildner mit Myrosinasen identifiziert wurden (Lenman et al. 1990; Falk et al. 1995; Geshi and Brandt 1998). Die biologische Bedeutung der MBPs ist bisher jedoch ungeklärt. Für die Faltung, die katalytische Aktivität und die Lokalisation der Myrosinasen in *B. napus* scheinen sie nicht ausschlaggebend zu sein. Ebenso konnte eine Beeinflussung des Wachstums der pathogenen Pilze *Alternaria brassicae* (Pleosporaceae) und *Lepthosphaeria maculans* (Phaeosphaeriaceae), des Fraßverhaltens des gewelltstreifigen Kohlerdflohs *Phyllotreta undulata* (Coleoptera: Chrysomelidae) und der Wachstumsrate des Mehlkäfers *Tenebrio molitor* (Coleoptera: Tenebrionidae) sowie der Hydrolyse von 4-Hydroxybenzylglucosinolat in *B. napus* nicht nachgewiesen werden (Eriksson et al. 2002).

Auf Grund phylogenetischer Untersuchungen wäre eine Funktion der Lektin-ähnlichen Jacalindomänen in der Interaktion von NSPs mit Myrosinasen denkbar (Burow et al. 2009). Eine Abstammung der Lektin-ähnlichen Jacalindomänen des AtNSP1, AtNSP2 und AtNSP3 sowie der zweiten Lektin-ähnlichen Jacalindomäne des AtNSP4 von den putativen MBP-Sequenzen MBP1 und MBP2 aus *A. thaliana* konnte nachgewiesen werden (Burow et al. 2009). Die erste Lektin-ähnliche Jacalindomäne des AtNSP4 scheint einen anderen Ursprung zu besitzen (Burow et al. 2009). Die Tatsache, dass AtNSP5 keine Lektin-ähnliche Jacalindomäne besitzt, eine Funktionalität als NSP jedoch biochemisch nachgewiesen werden konnte, deutet darauf hin, dass die Lektin-ähnlichen Jacalindomänen nicht essentiell für die NSP-Aktivität sind. Stattdessen könnte ihre Funktion in der gewebespezifischen Lokalisation und Regulation der betreffenden Nitril-spezifisierenden Proteine AtNSP1 – AtNSP4 liegen (Burow et al. 2009). Eine Bestätigung hierfür steht jedoch aus. Das Vorhandensein der Lektin-ähnlichen Jacalindomänen innerhalb der Sequenzen von AtNSP1 – AtNSP4 scheint daher, evolutiv betrachtet, den abgeleiteten Status der Proteine darzustellen.

Als Vorläuferprotein der spezifizierenden Proteine werden At3g07720 aus *A. thaliana* bzw. homologe Sequenzen anderer Spezies vermutet. Die Funktion von At3g07720 ist unbekannt. Das Protein stellt ein Homolog der in NSP-Assays aktiven spezifizierenden Proteine AtNSP1 - AtNSP5 dar. Trotz hoher Sequenzidentitäten konnte für At3g07720 keine Aktivität als spezifizierendes Protein nachgewiesen werden. Die basale Stellung von At3g07720 und homologer Sequenzen festigt sich durch die Tatsache, dass sich At3g07720 in phylogenetischen Analysen Homologen aus Spezies zuordnet, die keine Glucosinolate enthalten (Burow et al. 2009).

1.6 Funktionen freigesetzter Glucosinolat-Hydrolyseprodukte und ihre Rolle im „evolutiven Wettrennen“

Strukturbedingt unterscheiden sich die Glucosinolat-Hydrolyseprodukte sowohl in ihrer Flüchtigkeit als auch in ihrer biologischen Aktivität. Es konnte gezeigt werden, dass isolierte Isothiocyanate eine mit synthetischen Insektiziden vergleichbare toxische Wirkung aufweisen (Lichtenstein et al. 1962). Sowohl in der Gasphase als auch durch Kontakt wirken Isothiocyanate toxisch auf eine Vielzahl generalistischer Insekten sowie Pilze und Bakterien. Bei einigen Insekten treten zudem toxische Effekte nach Resorption der Isothiocyanate im Insektendarm auf (Wittstock et al. 2003). In Bioassays unter Verwendung intakter Pflanzen konnte gezeigt werden, dass sich Insekten auf Pflanzen mit geringerem Glucosinolatgehalt besser entwickeln (Li et al. 2000; Müller et al. 2010). Über die biologische Aktivität der alternativen Glucosinolat-Hydrolyseprodukte ist hingegen wenig bekannt. Zumindest die einfachen Nitrile scheinen im Vergleich zu Isothiocyanaten weniger giftig zu sein (Wittstock et al. 2003). So entwickelten sich Larven des Generalisten *Spodoptera littoralis* (Lepidoptera: Noctuidae) auf Nitril-produzierenden Linien von *A. thaliana* schneller als auf Isothiocyanat-produzierenden Linien (Burow et al., 2006a).

Dies wirft die Frage auf, warum Pflanzen in die Synthese der spezifizierenden Proteine investieren, wenn die resultierenden Produkte weniger wirksam sind. Als mögliche Antwort käme eine Rolle der alternativen Produkte in der indirekten pflanzlichen Abwehr in Betracht. Im Gegensatz zur direkten Abwehr, bei der giftige oder abschreckende Substanzen das Ausmaß der Herbivorie begrenzen, wirken bei der indirekten Abwehr z.B. flüchtige Pflanzenstoffe attraktiv auf Fraßfeinde eines Herbivoren. Direkte und indirekte Abwehr können im Rahmen der schrittweisen evolutiven Anpassung von Pflanzen und Herbivoren entstehen. Das Konzept der Coevolution wurde 1964 von Ehrlich und Raven am Beispiel von Schmetterlingen und ihren Wirtspflanzen vorgestellt. Dieses Konzept besagt, dass die Evolution sekundärer Pflanzenstoffe eine schrittweise evolutive Anpassung von Herbivoren und phytopathogenen Organismen nach sich ziehen kann. Diese Anpassung wirkt sich dann wieder auf die Evolution der sekundären Pflanzenstoffe aus. Die pflanzliche Anpassung kann ebenso wie der Wirtspflanzenwechsel eines Herbivoren die Eroberung einer neuen ökologischen Nische bedeuten und somit die Entstehung neuer Arten hervorrufen (Ehrlich and Raven 1964).

Ein interessantes Beispiel hierfür liefert der Spezialist *P. rapae*, welcher zur Familie der Pieridae gehört. Die Pieridae spalten sich in drei Unterfamilien, die Dismorphiinae, die Pierinae und die Coliadinae auf. Nur die Mitglieder der Pierinae ernähren sich von Glucosinolat-haltigen Brassicales während sich die Mitglieder der Dismorphiinae und fast alle Mitglieder der Coliadinae von Pflanzen der Ordnung Fabales ernähren (Braby and Trueman 2006). Ermöglicht wird die Lebensweise der Pierinae durch das im Darm von *P. rapae* Larven erstmals identifizierte insektoide NSP (PrNSP, AAR84202 (Wittstock et al. 2004)), dessen Aktivität später auch in anderen Vertretern der Pierinae nachgewiesen werden konnte. Das larvale NSP ermöglicht die Entschärfung des Glucosinolat-Myrosinase-Systems durch Umlenkung des Glucosinolatabbaus zu einfachen Nitrilen im Darm der Larven. Sie scheiden die Nitrile mit oder ohne weitere Metabolisierung mit ihrem Kot aus und können so

die toxische Wirkung der Isothiocyanate umgehen. Das insektoide NSP kann als eine Schlüsselinnovation bezeichnet werden, in deren Folge sich die Artenvielfalt der Pierinae vergrößerte (Wheat et al. 2007).

In Zusammenstellung dieser Erkenntnis und der Analyse bekannter Sequenzdaten sowie der Datierung der Familie der Pieridae durch Fossilienfunde konnte die Abspaltung der Pierinae ausgehend von der Unterfamilie der Coliadinae und der damit einhergehende Wirtspflanzenwechsel von den Fabales hin zu Glucosinolat-haltigen Brassicales auf ~ 85 Mill. Jahre eingeordnet werden. Übereinstimmend konnte das Auftreten der Ordnung Brassicales an Hand fossiler Daten in eine Zeitspanne von 90 Mill. Jahren - 85 Mill. Jahren einsortiert werden (Abbildung 5) (Wheat et al. 2007). Die Entstehung der Brassicales ist hierbei mit der Entstehung des Glucosinolat-Myrosinase-Systems gleichzusetzen. Vor ~ 71,3 Mill. Jahren entstand die Kerngruppe der Ordnung Brassicales (Abbildung 5) (Beilstein et al. 2010). Sie umfasst die Familien Brassicaceae, Capparaceae und Cleomaceae (Beilstein et al. 2010). Aus der Familie der Cleomaceae ging vor ~ 64,5 Mill. Jahren die Familie Brassicaceae hervor (Abbildung 5) (Beilstein et al. 2010). Die Entstehung des Artenreichtums der Haupt-Pierinae-Arten wurde ebenfalls auf ~ 65 Mill. Jahre geschätzt und korreliert somit mit der Entstehung der Familie der Brassicaceae (Abbildung 5) (Beilstein et al. 2010). Die Entstehung der Familie der Brassicaceae wird zudem in Verbindung mit einer vor ~ 70 Mill. Jahren eingetretenen Genomverdopplung (β) gebracht (Abbildung 5) (Beilstein et al. 2010). Die Familie Brassicaceae hat sich dann vor ~ 54,3 Mill. Jahren durch die Aufspaltung in die Aethionemeae und die Kerngruppe der Familie Brassicaceae aufgetrennt (Abbildung 5) (Beilstein et al. 2010). Diese Kerngruppe umfasst alle Abstammungslinien der Brassicaceae mit Ausnahme der Aethionemeae und fächerte sich vor ~ 46,9 Mill. Jahren auf (Abbildung 5) (Beilstein et al. 2010). Vor 43,2 Mill. Jahren kam es zur Differenzierung der Gattungen *Arabidopsis* und *Brassica* und im weiteren evolutiven Verlauf zur Auftrennung der drei Abstammungslinien (I: vor ~ 35,6 Mill. Jahren; II: vor ~ 30,8 Mill. Jahren; III: vor ~ 35,5 Mill. Jahren) (Abbildung 5) (Beilstein et al. 2010). Die Auffächerung der Brassicaceae-Kerngruppe lässt sich zeitlich mit einer Genomverdopplung (α) vor ~ 40 Mill. Jahren (und älter) in Verbindung bringen (Franzke et al. 2011).

Das larvale NSP der Pierinae ermöglichte somit vermutlich den Wirtspflanzenwechsel der Familie der Pieridae (Wheat et al. 2007). Es könnte zur Diversifizierung der Brassicaceae beigetragen haben, für die aber auch die α -Genomduplikation von großer Bedeutung war. Die möglichen Auswirkungen dieser Diversifizierung auf das Glucosinolat-Myrosinase-System könnten sich in einer Erweiterung des Glucosinolat-Hydrolyseproduktspektrums zeigen. Durch die Bildung zusätzlicher Hydrolyseprodukte könnte der Übergang von der direkten zur indirekten Abwehr ermöglicht worden sein.

Das folgende Beispiel zeigt deutlich, dass eine Veränderung des Spektrums der Glucosinolat-Hydrolyseprodukte innerhalb einer Pflanzenart diese auf Grund der verminderten Freisetzung von Isothiocyanaten einerseits anfälliger für generalistisch lebende Herbivoren macht, sie aber gegen einen Spezialisten verteidigen kann. Über Chemorezeptoren wahrgenommene Glucosinolate in der Blattoberfläche stimulieren Weibchen des Spezialisten *P. rapae* zur Eiablage (Renwick et al. 1992) und animieren ihre

Larven zur dortigen Nahrungsaufnahme während ihrer Entwicklung (Hopkins et al. 2009). Wird dann bei der Nahrungsaufnahme das pflanzliche Gewebe verletzt, initiiert dies in *A. thaliana* die Transkription und Translation des AtNSP1, welches den Glucosinolat-Abbau zu einfachen Nitrilen unterstützt (Burow et al. 2009). Da Isothiocyanate auch über größere Entfernungen von Spezialisten wahrgenommen werden können und diese anlocken, könnte die verringerte Menge an Isothiocyanaten die Wahrnehmung der Pflanze durch weibliche *P. rapae* verringern. Zusätzlich führt die erhöhte Menge an einfachen Nitrilen zur Anlockung der auf *P. rapae* parasitierenden Wespenart *Cotesia rubecula* (Hymenoptera: Braconidae) (Mumm et al. 2008). Spezifizierende Proteine könnten somit als Teil des Glucosinolat-Myrosinase-Systems innerhalb des beschriebenen evolutiven Wettrennens entstanden sein. Hinweise auf die Entstehung des heutigen Glucosinolat-Myrosinase-Systems liegen bisher jedoch fast ausschließlich auf Ebene der Glucosinolat-Biosynthese (1.2) und innerhalb der Myrosinasen (1.4) vor.

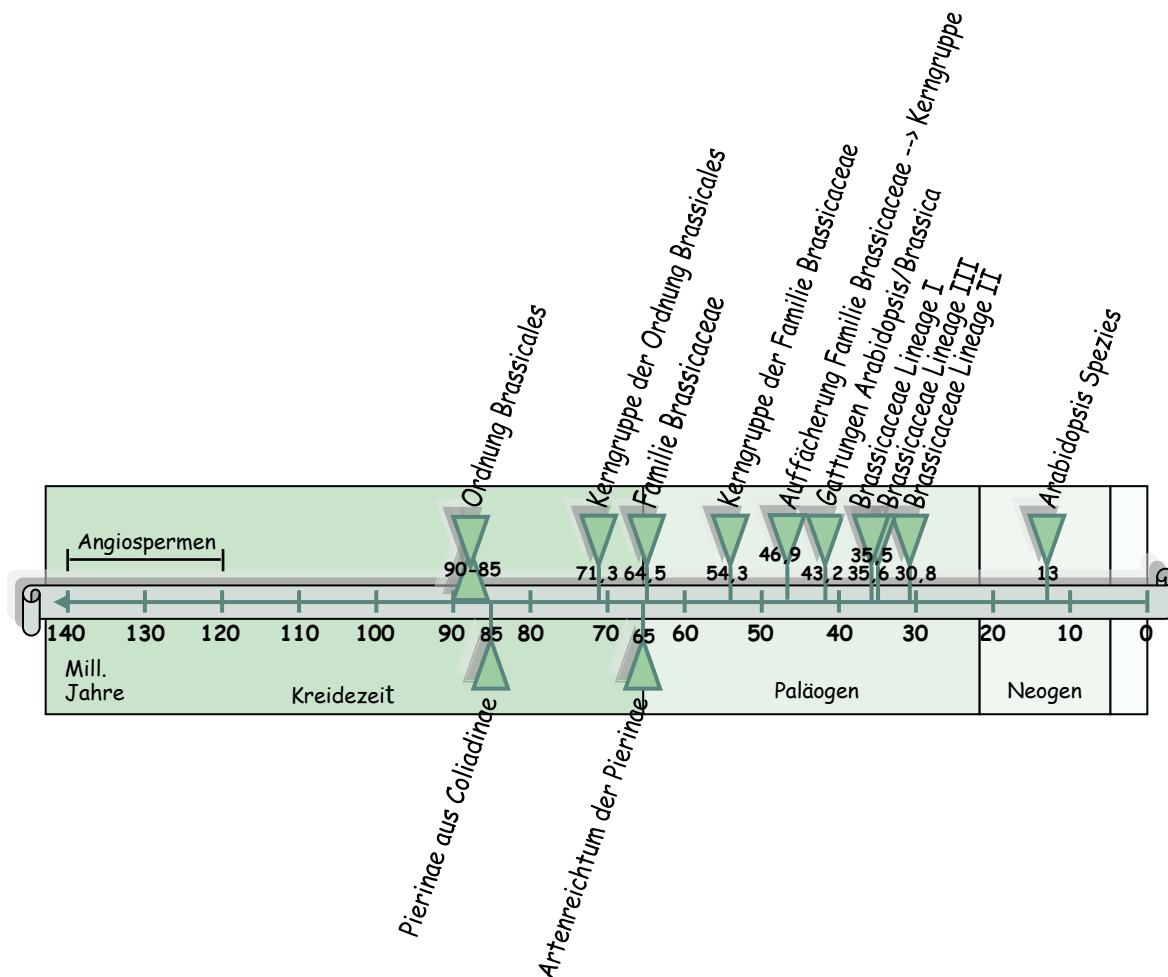


Abbildung 5: Evolutive Zeitleiste der Brassicales und Pierinae. Schematisch dargestellt ist die erdzeitgeschichtlich eingeordnete Entstehung der Ordnung Brassicales und ihrer Weiterentwicklung (oben) in Gegenüberstellung zur Entstehung und Entfaltung der Unterfamilie der Pierinae (unten) in Millionen Jahren. Quellen der verwendeten Daten sowie Einzelheiten der Zuordnung sind dem Text zu entnehmen.

1.7 Fragestellung und Zielsetzung der vorliegenden Arbeit

Alle bisher charakterisierten spezifizierenden Proteine (Tabelle 1) gruppieren sich in phylogenetischen Analysen innerhalb eines Zweiges getrennt von ähnlichen Sequenzen aus Pflanzen, die keine Glucosinolate enthalten. Als putativer Vorgänger erscheint At3g07720 an der Basis dieses Zweiges (Abbildung 6). Die Familie der spezifizierenden Proteine spaltet sich dann in NSPs, ESPs und TFPs auf. Die Stellung der ESPs, umgeben von den NSPs, innerhalb dieses Clusters spezifizierender Proteine legt den Rückschluss nahe, dass die Fähigkeit, Epithionitrile zu formen, eine hinzugewonnene Funktion ist, die sich auf der Basis der Fähigkeit zur Bildung einfacher Nitrile entwickelt hat (Burow et al. 2009). Dies wird dadurch unterstützt, dass ESPs je nach Substrat auch einfache Nitrile bilden können.

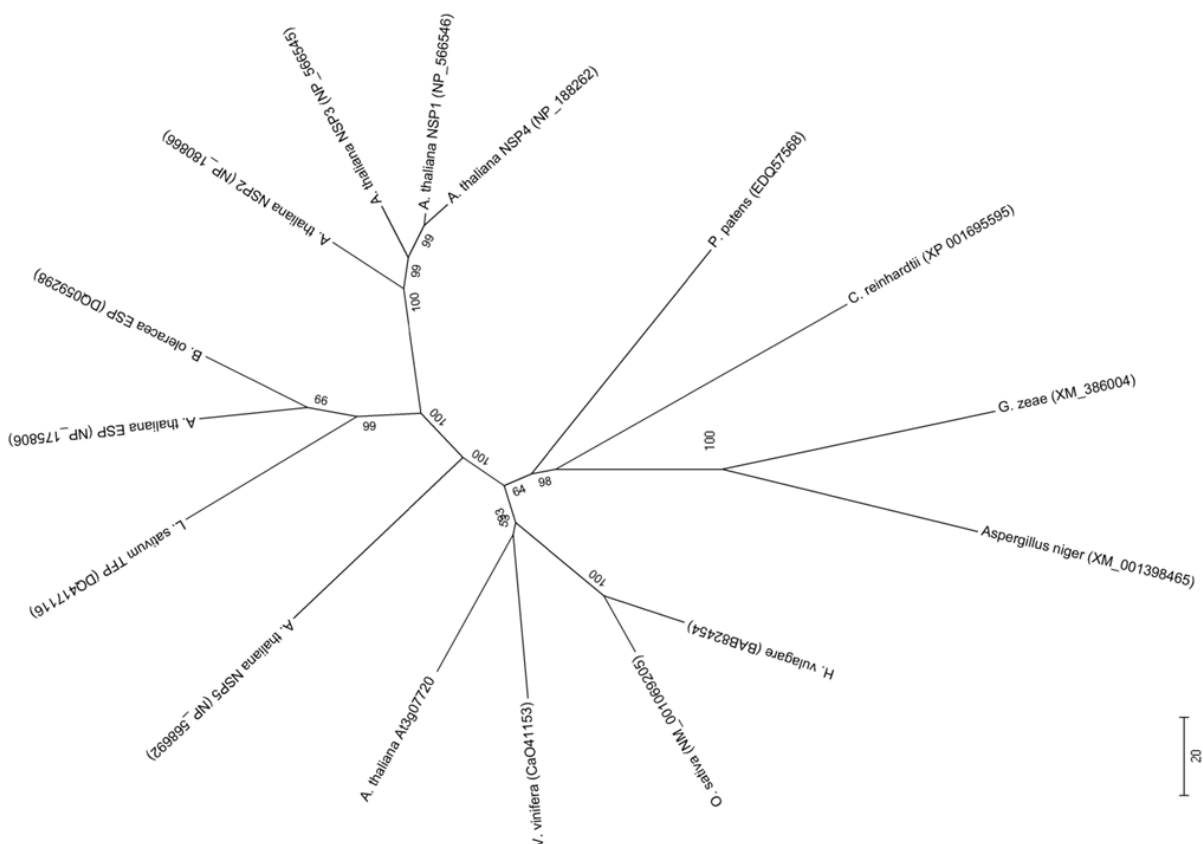


Abbildung 6: Phylogenie der Kelchdomänen von spezifizierenden Proteinen und ihren Homologen. Die Aminosäuresequenzen der Kelchdomänen von ESP Homologen höherer Pflanzen, einer Alge, einem Moos und zwei Pilzen wurden für die Rekonstruktion des Neighbor-joining-Baumes nach Burow et al. 2009 verwendet. Die Berechnung erfolgte unter Angabe einer 1000fachen Bootstrap-Wiederholung und wurde erzeugt unter Verwendung von Sequenzen der Spezies *A. thaliana*, *B. oleracea*, *Vitis vinifera*, *Chlamydomonas reinhardtii*, *Aspergillus niger*, *Physcomitrella patens*, *Hordeum vulgare* und *Oryza sativa*. Astlängen geben die geschätzten evolutionen Distanzen wieder. Die Maßstabsleiste zeigt die Länge für die angegebene Anzahl an Substitutionen pro Position.

Zielsetzung der vorliegenden Arbeit war die Klärung der von Burow et al. 2009 aufgestellten Hypothese einer Rolle der NSPs als Vorläufer aller spezifizierenden Proteine (Burow et al. 2009) und der Frage, ob At3g07720 und homologe Sequenzen die Basis einer Entwicklung zu NSPs darstellten. Desweiteren sollte die evolutive Entstehung von TFPs und ESPs ergründet werden. Sind TFPs direkt aus NSPs oder aus ESPs hervorgegangen?

Da die geringe Anzahl vorhandener Sequenzen spezifizierender Proteine bisher lediglich rein spekulative Aussagen zuließen, sollte durch die Isolierung von cDNAs und die biochemische Charakterisierung der entsprechenden Proteine ein umfassendes Repertoire an charakterisierten spezifizierenden Proteinen akquiriert werden. Phylogenetische Untersuchungen der erhaltenen Sequenzdaten sollten anschließend die aufgeführten Annahmen und Fragen beleuchten, um die Evolution der spezifizierenden Proteine nachvollziehen zu können. Da frühere Untersuchungen Hinweise auf posttranskriptionale Regulation der Expression von spezifizierenden Proteinen geliefert hatten (Burow et al. 2007a; Burow et al. 2007b), wurde zusätzlich eines der identifizierten spezifizierenden Proteine, ein aus *Thlaspi arvense* (*T. arvense*) isoliertes TFP, im Hinblick auf sein Expressionsmuster genauer untersucht.

2. ERGEBNISSE

2.1 Phytochemisches Screening zur Auswahl geeigneter Glucosinolat-haltiger Spezies

Frühere Arbeiten (Burow und Wittstock, persönliche Mitteilung) haben gezeigt, dass die Fähigkeit, andere Glucosinolat-Hydrolyseprodukte als Isothiocyanate zu bilden, nicht grundsätzlich bei allen Glucosinolat-haltigen Arten der Ordnung Brassicales auftritt. Für die Untersuchungen der vorliegenden Arbeit sollten Arten identifiziert werden, welche bei Gewebeverletzung Epithionitrile, organische Thiocyanate oder einfache Nitrile produzieren und daher mit hoher Wahrscheinlichkeit auch spezifizierende Proteine exprimieren. Diese Auswahl wurde auf der Basis eines phytochemischen Screenings getroffen. Hierfür wurden mittels HPLC-Analytik die Glucosinolatprofile von pflanzlichen Geweben ausgewählter Spezies bestimmt (Tabelle 2). Die Ergebnisse wurden durch die GC-MS-Analyse von Dichlormethanextrakten aus Homogenaten der jeweiligen Pflanzengewebe ergänzt, welche die Identifizierung der Glucosinolat-Hydrolyseprodukte ermöglichte (Tabelle 2). Im Rahmen der vorliegenden Arbeit wurden bereits vorliegende Daten (Burow und Wittstock, persönliche Mitteilung) durch die Untersuchung weiterer Spezies und Pflanzengewebe erweitert. In Tabelle 2 sind die im Rahmen dieser Arbeit erhaltenen Resultate zusammengetragen. Eine Übersicht des vollständigen phytochemischen Screenings ist dem Anhang zu entnehmen (Tabelle S 1).

Insgesamt lagen mit diesen und früheren Untersuchungen (Burow und Wittstock, persönliche Mitteilung) Daten für 30 Spezies der Ordnung Brassicales vor. Dabei ergab sich eine Auswahl von sechs Familien (Brassicaceae, Cleomaceae, Limnanthaceae, Moringaceae, Resedaceae, Tropaeolaceae) dieser Ordnung. Von Isothiocyanaten abweichende Glucosinolat-Hydrolyseprodukte konnten in 15 Spezies innerhalb der Familien Brassicaceae und Tropaeolaceae (*Tropaeolum minus*) nachgewiesen werden. In *Tropaeolum minus* (*T. minus*) waren einfache Nitrile vorhanden, konnten jedoch nur in sehr geringen Mengen detektiert werden. Zur besseren Übersicht sind diese 15 Spezies in Tabelle 3 gelistet. Sie wurden für die Klonierung spezifizierender Proteine eingesetzt mit Ausnahme von *Alyssum alpestre* (*A. alpestre*) und *T. minus*. Für diese Spezies stand nicht genug Samenmaterial zur Verfügung.

Tabelle 2: Glucosinolate und Glucosinolat-Hydrolyseprodukte in Spezies der Ordnung Brassicales. Glucosinolate wurden mittels HPLC-Analyse als Desulfoglucosinolate bestimmt. Glucosinolat-Hydrolyseprodukte wurden mittels GC-MS-Analyse in Dichlormethanextrakten von Homogenaten des Pflanzenmaterials identifiziert. Aufgeführt sind die im Rahmen der vorliegenden Arbeit erhaltenen Ergebnisse. Daten für weitere Spezies siehe Tabelle S 1. In der dritten Spalte der Tabelle sind unter „Glucosinolate“ die Seitenketten der identifizierten Glucosinolate aufgeführt (3MSOP = 3-Methylsulfinylpropyl, 3MTP = 3-Methylthiopropyl, 3OHP = 3-Hydroxypropyl, 4MSOB = 4-Methylsulfinylbutyl, 4MTB = 4-Methylthiobutyl, 5MSOP = 5-Methylsulfinylpentyl, 6MSOH = 6-Methylsulfinylhexyl, 6MTH = 6-Methylthiohexyl, 7MSOH = 7-Methylsulfinylheptyl, 7MTH = 7-Methylthioheptyl, 8MSOO = 8-Methylsulfinyloctyl, 8MTO = 8-Methylthiooctyl, 9MSON = 9-methylsulfinylnonyl, 10MSOD = 10-Methylsulfinyldecyl, 11MSON = 11-Methylsulfinylundecyl, 2-OH-3-butenyl = 2-Hydroxy-3-butenyl, 2-OH-2-Phenylethyl = 2-Hydroxy-2-Phenylethyl, 1MOI3M = 1-Methoxyindol-3-ylmethyl, 1,4-DiMOI3M = 1,4-Dimethoxy-indol-3-yl-methyl, 4MOI3M = 4-Methoxyindol-3-ylmethyl, 4OHI3M = 4-Hydroxy-indol-3-yl-methyl, I3M = Indol-3-yl-methyl).

Spezies	Organ	Glucosinolate	Hydrolyseprodukte
<i>Alliaria petiolata</i> (Sanct Joseph)	Keimlingswurzeln	Allyl, 4MOI3M, 4OHI3M, I3M, Benzyl,	Thiocyanat aus Allylglucosinolat, Epithionitril aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat, Nitril aus Benzylglucosinolat
	Keimblätter	Allyl, 4MOI3M, 4OHI3M, Benzyl	Thiocyanat aus Allylglucosinolat, Epithionitril aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat, Nitril aus Benzylglucosinolat
<i>Alyssum alpestre</i>	Blätter	3MTP, 3MSOP, 4MOI3M, 4MTB, 4MSOB, (unident. Alkyl),	Isothiocyanat aus 3-Methylthiopropylglucosinolat, Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat, Nitril aus 3-Methylthiopropyl
<i>Camelina microcarpa</i>	Blüten	3MSOP	Nicht identifizierbar
	Schoten	3MSOP, 10MSOD, 9MSON, 3MTP, I3M	Nicht identifizierbar
<i>Capsella rubella</i>	Blätter	10MSOD, 9MSON	Nicht identifizierbar
	Keimlinge	10MSOD, 9MSON	Nicht identifizierbar
	Samen	10MSOD, 9MSON	Nicht getestet
	Blüten		Isothiocyanat aus Allylglucosinolat
<i>Cardamine hirsuta</i>	Blätter	Benzyl, 3-Butenyl, I3M, 4-Pentenyl, 4MSOB, Phenylethyl, 4MOI3M, 4OHI3M, pOHB,	Nitril aus Benzylglucosinolat, Epithionitril aus 3-Butenylglucosinolat (unident. Indolderivat), Isothiocyanat aus Benzylglucosinolat, Nitril aus Phenylethylglucosinolat
<i>Cardamine impatiens</i>	Blätter	3-Butenyl, I3M, I-2-OH-3-Butenyl, (S)-2-OH-3-Butenyl, Benzyl, 4-Pentenyl, (unident. Alkyl)	Epithionitril aus 3-Butenylglucosinolat, Nitril aus Indol-3-yl-methylglucosinolat, Epithionitril aus (R/S)-2-Hydroxy-3-butenylglucosinolat,

<i>Cleome spinosa</i>	Blätter	3OHP, 4OHI3M, I3M	Nicht identifizierbar
	Samen	3OHP, Allyl, 4OHI3M, I3M	Nicht getestet
	Blüten	3OHP, 4OHI3M, , I3M, 1MOI3M	Methyl-ITC, Methylpropyl-ITC, Methylbutyl-ITC
<i>Diploaxis muralis</i>	Blätter	Nicht getestet	Isothiocyanat aus 4-Mercaptobutylglucosinolat
<i>Draba aurea</i>	Blätter	3-Butenyl, 4OHI3M, Allyl, 4-Pentenyl, 5MSOP, 9MSON, 4MSOB, 4MOI3M, I3M, Phenylethyl	Epithionitril aus 3-Butenylglucosinolat
<i>Draba cinerea</i>	Blätter	3MTP, 3MSOP, 4MTB, Methylethyl, (unident. Sulfonyl), 4MSOB, 9MSON, 8MSOO, 6MSOH, 6MTH oder 10MSOD?, 5MSOP, 7MSOH, 7MTH, 8MTO	Isothiocyanat aus 3-Methylthiopropylglucosinolat, Isothiocyanat aus 4-Methylthiopropylglucosinolat, Nitril aus 3-Methylthiopropylglucosinolat, Nitril aus 4-Methylthiopropylglucosinolat, Isothiocyanat aus Methylethylglucosinolat
<i>Draba lanceolata</i>	Blätter	I-2-OH-3-Butenyl, 3-Butenyl, 4MSOB, 5MSOP, 7MSOH, 9MSON, 8MSOO. Allyl,	Nicht getestet
<i>Moringa oleifera</i>	Blätter	pOHB, Benzyl, Allyl	Isothiocyanat aus Benzylglucosinolat
	Samen	Rhamnopyranosyloxybenzyl, pOHB	Nicht getestet
<i>Schouwia purpurea</i>	Blätter	Nicht getestet	Epithionitril aus 3-Butenylglucosinolat,
	Keimlinge	(S)-2-OH-3-Butenyl, 1MOI3M, 3-Butenyl, 4MOI3M, I3M, 3OHP, I-2-OH-3-Butenyl	Nitril aus Indol-3-yl-methylglucosinolat, Epithionitril aus (R/S)-2-Hydroxy-3-butenylglucosinolat, Epithionitril aus 3-Butenylglucosinolat, Nitril aus Indol-3-yl-methylglucosinolat
<i>Thlaspi arvense</i>	Blätter	Allyl, Benzyl, 3MSOP, 4MOI3M	Isothiocyanat aus Allylglucosinolat
			Epithionitril aus Allylglucosinolat,
	Keimlinge	Allyl, 1MOI3M, 4MOI3M, Benzyl, I3M, 3MSOP, 4OHI3M	Thiocyanat aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat
	Samen	Allyl, Benzyl, I3M	Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat
	Wurzeln	Allyl, 1MOI3M, 4MOI3M, Benzyl, I3M	Epithionitril aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat, Isothiocyanat aus Benzylglucosinolat

<i>Tropaeolum minus</i>	Blüten	Allyl, Benzyl, 4MOI3M, 4OHI3M	Nitril aus Benzylglucosinolat Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat, Nitril aus Allylglucosinolat
	Schoten	Allyl, 3MSOP, Benzyl, 4OHI3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat
	Stängel	Allyl, 4MOI3M, Benzyl, 4OHI3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat,
	Blüten	Benzyl, pOHB	Isothiocyanat aus Benzylglucosinolat, Nitril aus Benzylglucosinolat

Tabelle 3: Spezies der Ordnung Brassicales, die alternative Glucosinolat-Hydrolyseprodukte bilden. Auflistung aller Spezies und Zugehörigkeiten zum jeweiligen Tribus derjenigen Mitglieder der Ordnung Brassicales, in deren Homogenaten andere Glucosinolat-Hydrolyseprodukte als Isothiocyanate nachgewiesen wurden. Das Vorhandensein von einfachen Nitrilen, Epithionitrilen oder organischen Thiocyanaten wurde durch + gekennzeichnet.

Spezies	Tribus	Epithionitril	Nitril	Thiocyanat
<i>Alliaria petiolata</i>	Thlaspidaceae	+	+	+
<i>Alyssum alpestre</i>	Alysseae		+	
<i>Cardamine hirsuta</i>	Cardamineae	+	+	
<i>Cardamine impatiens</i>	Cardamineae	+	+	
<i>Diplotaxis muralis</i>	Brassicaceae		+	
<i>Draba aurea</i>	Arabideae	+		
<i>Draba cinerea</i>	Arabideae		+	
<i>Draba lanceolata</i>	Arabideae	+	+	
<i>Draba muralis</i>	Arabideae		+	
<i>Erysimum hieraciifolium</i>	Brassicaceae		+	
<i>Iberis amara</i>	Iberideae		+	
<i>Isatis tinctoria</i>	Isatideae		+	
<i>Schouwia purpurea</i>	Brassicaceae	+	+	
<i>Thlaspi arvense</i>	Thlaspidaceae	+		+
<i>Tropaeolum minus</i>	Antirrhineae		+	

2.2 Wahl einer Klonierungsstrategie

Zur Berechnung eines aussagekräftigen phylogenetischen Stammbaumes ist eine ausreichend große Anzahl repräsentativer Sequenzen biochemisch charakterisierter Proteine notwendig. Die Anzahl bekannter Sequenzen (Tabelle 1) sollte im Rahmen dieser Arbeit erweitert werden, um an Hand des vergrößerten Datensatzes eine zutreffende Aussage über die phylogenetischen Beziehungen der spezifizierenden Proteine des Glucosinolat-Myrosinase-Systems zu erhalten.

Mit Hilfe multipler Sequenzvergleiche zwischen den bekannten Sequenzen spezifizierender Proteine (Tabelle 1) wurde zunächst geprüft, ob eine Strategie mittels Polymerase-Kettenreaktion (PCR) erfolgreich sein könnte. Durch den Einsatz degenerierter Oligonukleotide (Primer) sollte versucht werden, Sequenzvariationen der spezifizierenden Proteine zu berücksichtigen und trotzdem eine Differenzierung zwischen *NSP*- und *ESP / TFP*-Sequenzen während der PCR-Reaktion zu ermöglichen. Unter Verwendung des EBI Webservers für multiple Sequenzalignments ClustalW2 (Chenna et al. 2003) wurden mehrere Alignments auf Ebene der Aminosäuresequenzen erstellt. Dadurch konnten konservierte Regionen identifiziert werden, welche sich zur Generation von degenerierten Primern mit möglichst geringem Degenerationsgrad eigneten (Abbildung 7). Um den Grad der Degeneration gering zu halten, wurden zusätzlich zu den Aminosäuresequenzen auch die Nukleotidsequenzen in den ausgewählten Bereichen berücksichtigt. Die degenerierten Primer (Abbildung 7) wurden auf Basis dieses Vergleiches erstellt. Die genauen Primersequenzen sind Tabelle S 10 im Anhang zu entnehmen.

GTGGCGAGTTACACCAAATCAGCCCATC--GACAAACACCTTTACGTCT	592	CCACACCTCTCTTGTGTTAGGTGTCCGGATGGTGTCTAGTAGGATCGACCC	692
GTGGTGAGCTCACACCAAATCAGCCCATC--GACAAACACCTTTACGTCT	595	CCAAACCTCTCTTGTGTTAGGCGTCCGATGGTGTCAATTGGATCAAGCCT	695
GTGGCGAGTTAACTCCCAACATTTCCATC--GACAAAGACCTTTACGTCT	163	CCTGACGTCAAGGCCTTGGGCACCCGATGGTGTCCGTGGGAACCTAAGCT	263
GCGGCGAGTTAACACCAAACAAACATC--GACAAAGACCTTCTACGTCT	163	CCAACGTATCCTGCTTAGGCGTGCATGGTGGCCGTGGGAACCTAAGAT	263
GCGGCGAGCTGAAACACAGTTTACCTTC--GACAAACACCTTTACGTCT	163	CCTAGCCTCAGCTGCTTAGGTGTGCGCATGGTGGCCGTAGGAACCTAAGAT	263
GTGGTGAGCTTAAGCCAA--CGATCCATATTGATAACGATCTTTACGTTT	157	CC--TTTCCCTGTTTGGTGTCTCGATGGTAACGATCGGTTCAACGAT	254
* * * * *			
TTGACCTCGAGACCCGAGCTTGGTCCATTTCTCCAGCCACCGGAGACGTT	642	CTATGCTCTTTGGAGGCCGAGAGCCTTACGCCAATAACAACGGTTTCTACT	742
TTGACCTCGAGACCCGAGCTTGGTCCATTTCTCCAGCCACCGGAGACGTT	645	CTATGCTCTTTGGAGGCCGAGAGCCTTACGCCAATAACAACGGTTTCTACT	745
TTGACTTCAACACTCACAATTGGTCAATCTCTCCGTCCAAGGAGTAGCC	213	CTATCTATTCCGAGGCCGCGACGAGAATAAAAGTTTCGATGACTTTTATT	313
TTGACTTCAACACTCAAACTTGGTCAATCGCTCAACCCAAAGGAGACGCC	213	CTATATCTTTGGAGGCCGCGATGAGAACCCGAACCTTCGAAAACCTTCGCT	313
TTGACTTCAACAGCAACACTTGGTCCATCGCTGAGCCCAAAGGAGAGGCC	213	CTATATCTTTGGAGGCCGCGATGAGAACCCGAACCTTCGAAAACCTTCGCT	313
TCGATCTTGAGACTCAAGAAATGGTCGATAGCTCTCGCACCAGGAGAGCT	207	CTATGCTCTACGGTGGCCGTGACGATAAACGGAGATACAACGGTCTTCATT	304
* * * * *			
<i>Arabidopsis thaliana</i> NSP1	1	MAQKLEAKGGEMGDVWDDGVYENVRKVYVQQAQYGFVVKFEYVNGSQVVGDEHGKKTE	
<i>Arabidopsis thaliana</i> NSP2	1	MVQKVEARGGEIGDVWDDGAYDGVKRVYVQGEGDGFVVKFEYVNGSQEVVGDERGKCTL	
<i>Brassica oleracea</i> ESP	1	-----	
<i>Arabidopsis thaliana</i> ESP	1	-----	
<i>Lepidium sativum</i> TFP	1	-----	
<i>Arabidopsis thaliana</i> NSP5	1	-----	
<i>Arabidopsis thaliana</i> NSP1	61	LGVEEFEIDADDYIVYVEGYREKVNMTS-EMITFLSIKTFKGTSHPIEKRPVKFVVLH	
<i>Arabidopsis thaliana</i> NSP2	61	LGAEFEVDPPDYIVYVEGYHEKVFGVTKEIISTLTFTKYKGTSPPPFIVSGTKFVLQ	
<i>Brassica oleracea</i> ESP	1	-----	
<i>Arabidopsis thaliana</i> ESP	1	-----	
<i>Lepidium sativum</i> TFP	1	-----	
<i>Arabidopsis thaliana</i> NSP5	1	-----	
<i>Arabidopsis thaliana</i> NSP1	120	GGKIVGFHGRSTDVLHSLGAYVLSSTIKLGGKWKVQKQ-EGPGLRCSHGIAVGVNKL	
<i>Arabidopsis thaliana</i> NSP2	121	GGKIVGFHGRSTDVLHSLGAYISSPATPKLRGKWKVQKQ-EGPGLRCSHGIAVGVNKL	
<i>Brassica oleracea</i> ESP	1	-----MAPSVQGEWIKVQKQGGTTPGPRSSHGIAVGVNKL	
<i>Arabidopsis thaliana</i> ESP	1	-----MAPTVQGEWIKVQKQGGTTPGPRSSHGIAVGVNKL	
<i>Lepidium sativum</i> TFP	1	-----MALTVQGEWIKVQKQGGTTPGPRSSHGIAVGVNKL	
<i>Arabidopsis thaliana</i> NSP5	1	-----MCPVENKWLKVGQKQ-AGPGARSSHAITVVGKLV	
<i>Arabidopsis thaliana</i> NSP1	179	YSFGGEFTFNQPIDKHLYVFDLEIRTWISIPATGDPVHISCLGVRMVSIGSTLYVFGGRD	
<i>Arabidopsis thaliana</i> NSP2	180	YSFGGEFTFNQPIDKHLYVFDLEIRTWISIPATGDPVHISCLGVRMVSIGSTLYVFGGRD	
<i>Brassica oleracea</i> ESP	36	YSFGGEFTFNISIDKDLVYFDNFHTWISISPKGVAPDVKALGVRMVSIGSTLYVFGGRD	
<i>Arabidopsis thaliana</i> ESP	36	YSFGGEFTFNKHIDKDLVYFDNFHTWISIAQPKGAPDVISCLGVRMVSIGSTLYVFGGRD	
<i>Lepidium sativum</i> TFP	36	YVFGGEFTFNQPIDKHLYVFDLEIRTWISIAEPKGEAPSVISCLGVRMVSIGSTLYVFGGRD	
<i>Arabidopsis thaliana</i> NSP5	34	YCFGGEFTFNIDNDLYVFDLEIRQEWISIPATGAPFPFVGVSMVITGSTLYVFGGRD	
<i>Arabidopsis thaliana</i> NSP1	239	ASRQYNGYSYSEDITTNWKLLTPVEE--GPTPRSFHSMADENNYYVFGGVV-----AT	
<i>Arabidopsis thaliana</i> NSP2	240	ASRQYNGYSYSEDITTNWKLLTPVEE--GPTPRSFHSMADENNYYVFGGVV-----AT	
<i>Brassica oleracea</i> ESP	96	ENRKEDEYSYDITVTNEWTKLTLIDQEGGPEARTYHSMASDENHYVYVFGGVSKGGTNKTP	
<i>Arabidopsis thaliana</i> ESP	96	ENRNFENERSYDITVTSEWTFLLKLDVVGGEARTFHSMAADENNHYVYVFGGVSKGGTMNTP	
<i>Lepidium sativum</i> TFP	96	ENRNYSDYSYDITVKKEWKFLLKLDVVGGEARTFHSMAADENNHYVYVFGGVSKGGVQSTP	
<i>Arabidopsis thaliana</i> NSP5	93	DKERYNGLHSYDITETNEWKLLAPVEE--GLPGRSYHSMAGDDRKYVYVFGGVV-----AK	
<i>Arabidopsis thaliana</i> NSP1	291	ARLNTLDSYNIYDKKWFHCSTPGD---SLTARGGAGLEVVGKLVVVV-----GFN	
<i>Arabidopsis thaliana</i> NSP2	292	VRLLTLDAYNIYDKKWFHCSTPGG---SCSVRGGAGLEVVGKLVVVV-----GFN	
<i>Brassica oleracea</i> ESP	156	FRFRTIEAYNIADGKWSQLPDPCGEQFPFRERRGGAGFVVVGKLVVVVGFATSPDPNGKN	
<i>Arabidopsis thaliana</i> ESP	156	TRFRTIEAYNIADGKWAQLPDPCD---NFEKRGAGFAVVVGKLVVVVGFATSIIVPGKGD	
<i>Lepidium sativum</i> TFP	156	FRFKSTIVYNIADGTWSQLPNPGP---DFEPRGGAGLAVDDKLVVVVCGFANS-TSGGIN	
<i>Arabidopsis thaliana</i> NSP5	145	GRVNTLHAYDVVVGKWEYPAAGE---ACKGRGAPGLVVVVEGRVWLV-----GFD	
<i>Arabidopsis thaliana</i> NSP1	339	GCEVDDVHYYPDQDKWTQVETFCVPRPSEERSVFAAATCKHIVIFGGFIAMDPLAHVGP	
<i>Arabidopsis thaliana</i> NSP2	340	GCEVDDVHCYDPAQDKWTQVETFCGKPCRSVFAAVVCKHIVIFGGFIAMDPRHGP	
<i>Brassica oleracea</i> ESP	216	DYSDQVQFYDPATQKWEVETFCDKPSERSVFAHVVGKYIIFGGFTWPDPKHHLGP	
<i>Arabidopsis thaliana</i> ESP	213	DYENSAVQFYDPASKKWTEVETFCAPKPSERSVFAHVVGKYIIFGGFTWPDPKHHLGP	
<i>Lepidium sativum</i> TFP	212	DYNSNKVQYDVLVSGKWIEVETFCVVKPSERSVFAVAVGKQIVIFGGFIAMDPRHGP	
<i>Arabidopsis thaliana</i> NSP5	193	GNELGDIHCFDLASEQKAVETFCDVPAERSVFAVAVGKYIIFGGFIAMDPRHGP	

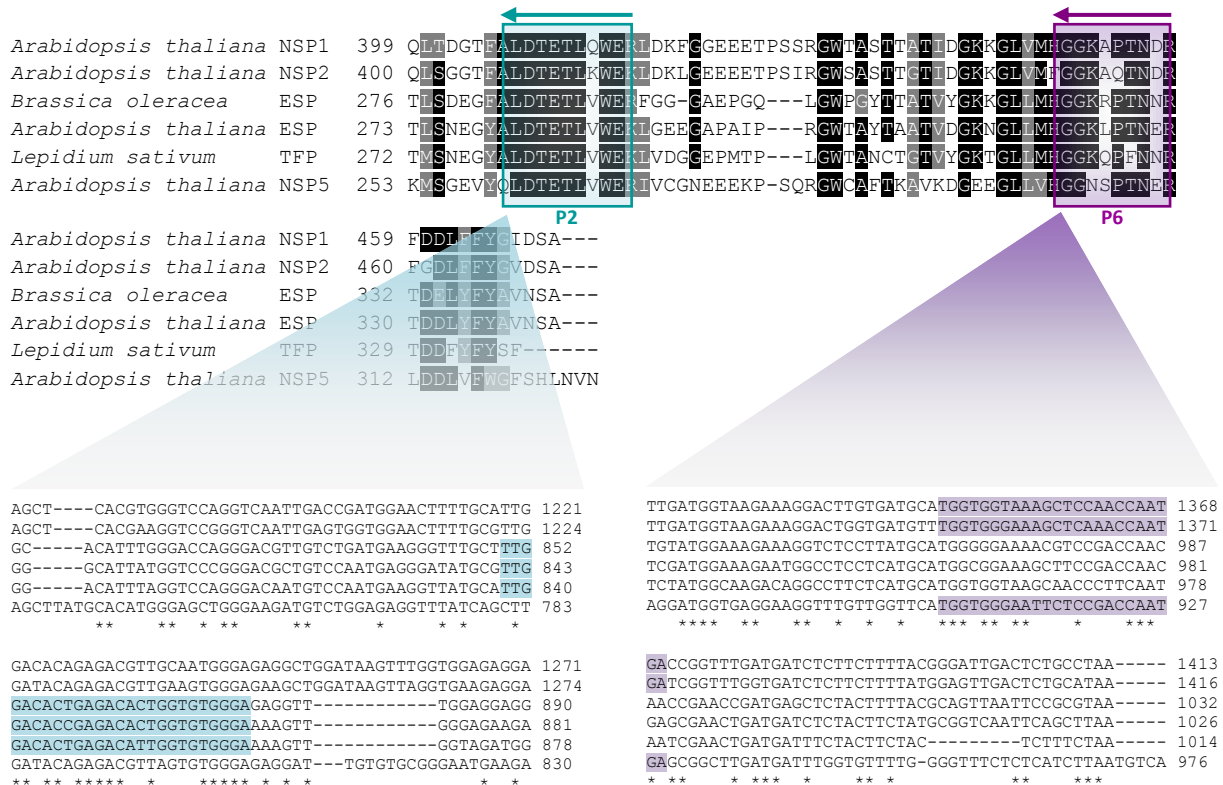


Abbildung 7: Auswahl von Positionen für degenerierte Primer auf der Basis eines Sequenzvergleiches bekannter spezifizierender Proteine: Alignment der Aminosäuresequenzen von AtESP, BoESP, AtNSP1, AtNSP5 und LsTFP. Erstellt wurde das Alignment mit dem EBI Webserver ClustalW2 und schattiert mittels des EMBnet Webserver BOXSCHADE 3.21. Weiße Buchstaben auf schwarzem Grund zeigen Aminosäuren einer Spalte, die in mindestens 80 % der Sequenzen identisch sind. Weiße Buchstaben auf grauem Grund zeigen Aminosäuren einer Spalte, die in mindestens 80 % der Sequenzen einer Gruppe mit biochemisch ähnlichen Eigenschaften zugeordnet werden können. Positionen degenerierter Primer sind auf den Aminosäuresequenzen durch Kästen mit Pfeilen darüber gekennzeichnet. Die Einschübe zeigen jeweils das Alignments der Nukleotidsequenzen in den Regionen hoher Aminosäuresequenzidentität. Die Primerpositionen sind farblich unterlegt. P1 und P2 (blau) für ESP- und TFP-Sequenzen. P5 und P6 (violet) für NSP-Sequenzen.

2.3 Identifizierung von cDNA Sequenzen putativer spezifizierender Proteine

Zur Klonierung putativer spezifizierender Proteine kam eine „*Degenerate Oligonucleotide Primer-PCR*“-Strategie (DOP-PCR) zum Einsatz. Hierfür wurde RNA aus Pflanzenmaterial isoliert und in cDNA transkribiert. Die cDNA wurde anschließend als Template für die DOP-PCR eingesetzt. Das Primerpaar P1 und P2 wurde mit dem Ziel verwendet *ESP*- und *TFP*-Sequenzen zu amplifizieren und das Primerpaar P5 und P6 mit dem Ziel *NSP*-Sequenzen zu amplifizieren. Die Größe der zu erwartenden DNA-Fragmente lag bei Verwendung des Primerpaares P1 und P2 bei 721 bis 733 Nukleotiden (entsprechend 240 bis 244 Aminosäuren) sowie bei Verwendung des Primerpaares P5 und P6 bei 670 bis 679 Nukleotiden (entsprechend 223 bis 226 Aminosäuren). PCR-Fragmente in der erwarteten sowie abweichenden Größe wurden für die in Tabelle 4 gelisteten Spezies und angegebene Primerpaare erhalten. Die Analyse der DOP-PCR-Ansätze ist in Abbildung 8 beispielhaft dargestellt.

Tabelle 4: Ergebnisse der DOP-PCR: Gelistet sind alle zur Isolation von RNA verwendeten Spezies und die in der DOP-PCR eingesetzten Primerpaare. Eine erfolgreiche DOP-PCR der eingesetzten Primerpaare ist durch + gekennzeichnet. Erhaltene DNA-Fragmente der erwarteten Größenordnung sind mit + gekennzeichnet. Abweichende Größen der PCR-Fragmente lagen zwischen 150 bis 400 Nukleotiden und sind ebenfalls mit + gekennzeichnet. n.i = nicht identifizierbar

Spezies	P1 und P2		P5 und P6	
	721bp - 733bp	150bp – 400bp	670bp – 679bp	150bp – 400bp
<i>Alliaria petiolata</i>	+	+	+	+
<i>Cardamine hirsuta</i>	+		+	
<i>Cardamine impatiens</i>	+		+	
<i>Diplotaxis muralis</i>	n.i		n.i	
<i>Draba aurea</i>	+		+	+
<i>Draba cinerea</i>	+		+	+
<i>Draba lanceolata</i>	+		+	+
<i>Draba muralis</i>	+			
<i>Erysimum hieraciifolium</i>	+			
<i>Iberis amara</i>	+			
<i>Isatis tinctoria</i>	+		+	
<i>Schouwia purpurea</i>	+		+	
<i>Thlaspi arvense</i>	P1 / Oligo (dT)-Primer siehe Text unterhalb			

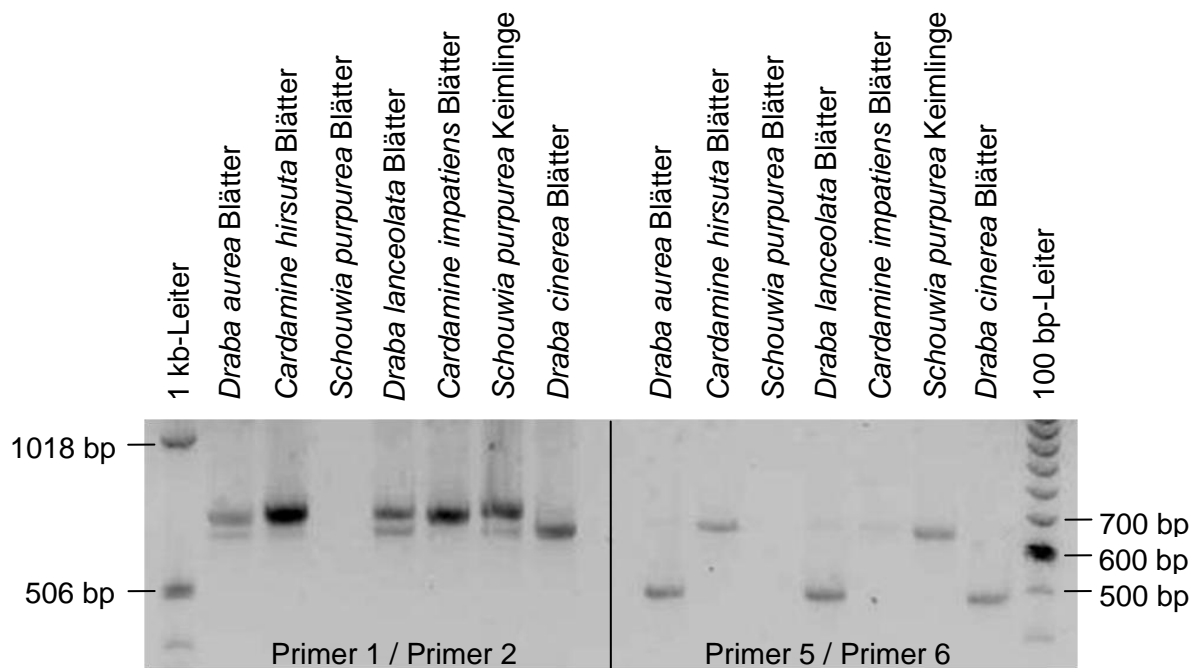


Abbildung 8: DOP-PCR an ausgewählten Spezies. cDNA der angegebenen Spezies wurde als Template in der DOP-PCR mit den angegebenen Primern eingesetzt. Anschließend wurden die PCR-Ansätze mittels Agarose-Gelelektrophorese analysiert und Banden unter Verwendung von Ethidiumbromid im UV-Licht detektiert. Gezeigt ist das Lichtbild eines Agarosegels. Zum Größenvergleich wurden links eine 1 kb-Leiter und rechts eine 100 bp-Leiter aufgetragen. Die Größenzuordnungen der Banden der Leitern sind angegeben.

Alle detektierbaren Banden der erwarteten Größenordnung sowie kleinere Banden mit 150 bis 450 Nukleotiden wurden aus dem Agarosegel ausgeschnitten und die DNA-Fragmente in *Escherichia coli* kloniert (pGEM®-T Easy Vektor, *E. coli*-Stamm XL1Blue). Eine nachfolgende DNA-Sequenzierung zeigte, dass insgesamt 17 DNA-Fragmente mit Ähnlichkeit zu cDNAs bekannter spezifizierender Proteine aus DNA-Fragmenten der erwarteten Größenordnung erhalten werden konnten (Tabelle 5). DNA-Fragmente der abweichenden Größenordnung führten nicht zu verwertbaren Ergebnissen. Ohne Ergebnis blieben DOP-PCR-Ansätze auf cDNA der Spezies *Diplotaxis muralis*.

Auf Basis der erhaltenen Sequenzen wurden genspezifische Primer für die anschließende Amplifikation der 5'- und 3'- Bereiche der cDNAs inklusive Start- und Stoppcodons mittels 5' "Rapid amplification of cDNA-ends-PCR" (RACE-PCR) und 3'RACE-PCR generiert (Tabelle S 10). Für die meisten DOP-PCR-Fragmente gelang es, die 5'RACE- und 3'RACE-PCR erfolgreich durchzuführen (Tabelle 5). Die erhaltenen RACE-Fragmente wurden kloniert und sequenziert.

Eine leicht abweichende Klonierungsstrategie kam für die Isolation eines spezifizierenden Proteins aus *T. arvense* zum Einsatz. Hier wurde das Ausgangs-DNA-Fragment unter Einsatz von P1 mit einem Oligo-(dT)-Primer (JCK Anchor d(T)18, P7, Tabelle S 10) gewonnen und beinhaltete bereits das Stoppcodon. Nur der 5'Bereich wurde mittels 5'RACE-PCR gewonnen.

Nach Zusammenstellung aller pro putativer cDNA erhaltenen Fragmente wurden an Hand dieser Sequenzdaten Primer erstellt, mit deren Hilfe in einer PCR-Reaktion unter Verwendung der cDNA als Template, die *Full-length* cDNAs amplifiziert werden konnten (Tabelle S 10). Die entsprechende PCR der *Full-length* cDNAs erfolgte nach Möglichkeit in jeweils zwei unabhängigen Ansätzen und bestätigte die Richtigkeit der im Vorfeld putativen offenen Leserahmen (ORF). Auf diese Weise konnten elf cDNAs von Start- bis Stoppcodon erhalten werden (Tabelle 5). Darüber hinaus konnten für einige cDNAs auch Teile der 5' und 3' untranslatierten Bereiche (UTR) amplifiziert werden. Nur unvollständige DNA-Fragmente konnten aus *Draba cinerea* (*D. cinerea*) und *Erysimum hieraciifolium* (*E. hieraciifolium*) isoliert werden (Tabelle 5), sie wurden in die weitere funktionelle Charakterisierung mit anschließender phylogenetischer Zuordnung nicht einbezogen. Weiterhin konnte eine Sequenz aus *Iberis amara* (*I. amara*) isoliert werden, welche entsprechend Abschnitt 2.4 als ESP eingeordnet werden konnte. Auf Grund ihrer hohen Aminosäureidentität von 97 % zu einer zuvor aus *Isatis tinctoria* (*I. tinctoria*) isolierten Sequenz wurde diese jedoch als Artefakt unbekannten Ursprungs gewertet und in die folgenden Untersuchungen nicht einbezogen

Tabelle 5: Auflistung aller mittels DOP- und RACE-PCR gewonnenen DNA-Fragmente: cDNAs der angegebenen Spezies wurden unter Verwendung jeweils der Primerkombination P1 mit P2 und P5 mit P6 zur DOP-PCR eingesetzt. Durch die Anzahl der gewonnenen DNA-Fragmente markierte Primerkombinationen führten zur erfolgreichen Amplifikation eines Fragmentes mit Ähnlichkeit zu cDNAs bekannter spezifizierender Proteine, Die erfolgreiche Amplifikation zugehöriger 3' und 5' Sequenzbereiche ist mit + angegeben. Eine gelungene Amplifikation der vollständigen ORFs ist durch + markiert.

Spezies	P1 + P2	3'Race	5'Race	ORF	P5 + P6	3'Race	5'Race	ORF
<i>Alliaria petiolata</i>	Fragment 1	+	+	+	Fragment 3	+	+	+
	Fragment 2	+						
<i>Cardamine hirsuta</i>	Fragment 1	+	+	+				
	Fragment 2	+						
<i>Cardamine impatiens</i>	Fragment 1	+	+	+				
<i>Draba aurea</i>	Fragment 1	+	+	+				
<i>Draba cinerea</i>	Fragment 1	+						
	Fragment 2							
<i>Draba lanceolata</i>	Fragment 1	+	+	+				
<i>Erysimum hieraciifolium</i>	Fragment 1							
<i>Isatis tinctoria</i>	Fragment 1	+	+	+	Fragment 3			
	Fragment 2	+	+	+				
<i>Schouwia purpurea</i>	Fragment 1	+	+	+	Fragment 2	+	+	+
<i>Thlaspi arvense</i>	Fragment 1	+	+	+				
P1 / Oligo (dT)-Primer								

2.4 Funktionelle Charakterisierung der identifizierten spezifizierenden Proteine

Zur funktionellen Charakterisierung wurden die erhaltenen cDNAs von Start- bis Stoppcodon mittels „*Uracil-Specific Excision Reagent*“-Klonierung (USER-Klonierung (Nour-Eldin et al. 2006)) in modifizierte Expressionsvektoren des pET-Expressionssystems überführt und im *E. coli*-Stamm BL21(DE3) pLysS zur Expression gebracht. Anschließend wurden bakterielle Rohextrakte oder die daraus aufgereinigten Proteine auf ihre Aktivität als spezifizierendes Protein untersucht. Dazu wurden Rohextrakte oder gereinigte Proteine mit Allyl- oder Benzylglucosinolat und Myrosinase inkubiert. Erhaltene Glucosinolat-Hydrolyseprodukte konnten im Folgenden mit Dichlormethan extrahiert und durch GC-MS-Analytik identifiziert werden. Eine funktionelle Zuordnung der spezifizierenden Proteine wurde auf Grund der im Assay entstandenen Glucosinolat-Hydrolyseprodukte vorgenommen. Drei spezifizierende Proteine, welche bei der Myrosinase-Reaktion ausschließlich die Bildung einfacher Nitrile aus Allylglucosinolat und Benzylglucosinolat katalysierten, wurden als NSPs eingestuft (Abbildung 9 A, B). Sechs spezifizierende Proteine, welche bei Allylglucosinolat-Hydrolyse Epithionitrile bildeten, wurden den ESPs zugeordnet (Abbildung 9 C, D). Zwei Proteine wurden den TFPs zugeordnet, da sie die Bildung organischer Thiocyanate katalysieren (Abbildung 9 E, F). Auffällig war hierbei, dass die im Rahmen dieser Arbeit identifizierten, TFPs aus *Alliaria petiolata* (*A. petiolata*) und *T. arvense* (JN244735 (Kuchernig et al. 2011)) organisches Thiocyanat aus Allylglucosinolat, aber nicht aus Benzylglucosinolat bildeten (Abbildung 9 C, D, E, F). Im Rahmen der vorliegenden Arbeit konnten also elf neue spezifizierende Proteine identifiziert und biochemisch charakterisiert werden. Zusammen mit den bereits bekannten acht spezifizierenden Proteinen (Tabelle 1) standen somit 19 charakterisierte spezifizierende Proteine (Tabelle 6) und die ähnliche Sequenz At3g07720 (Burow et al. 2009; Kissen and Bones 2009) für phylogenetische Analysen zur Verfügung.

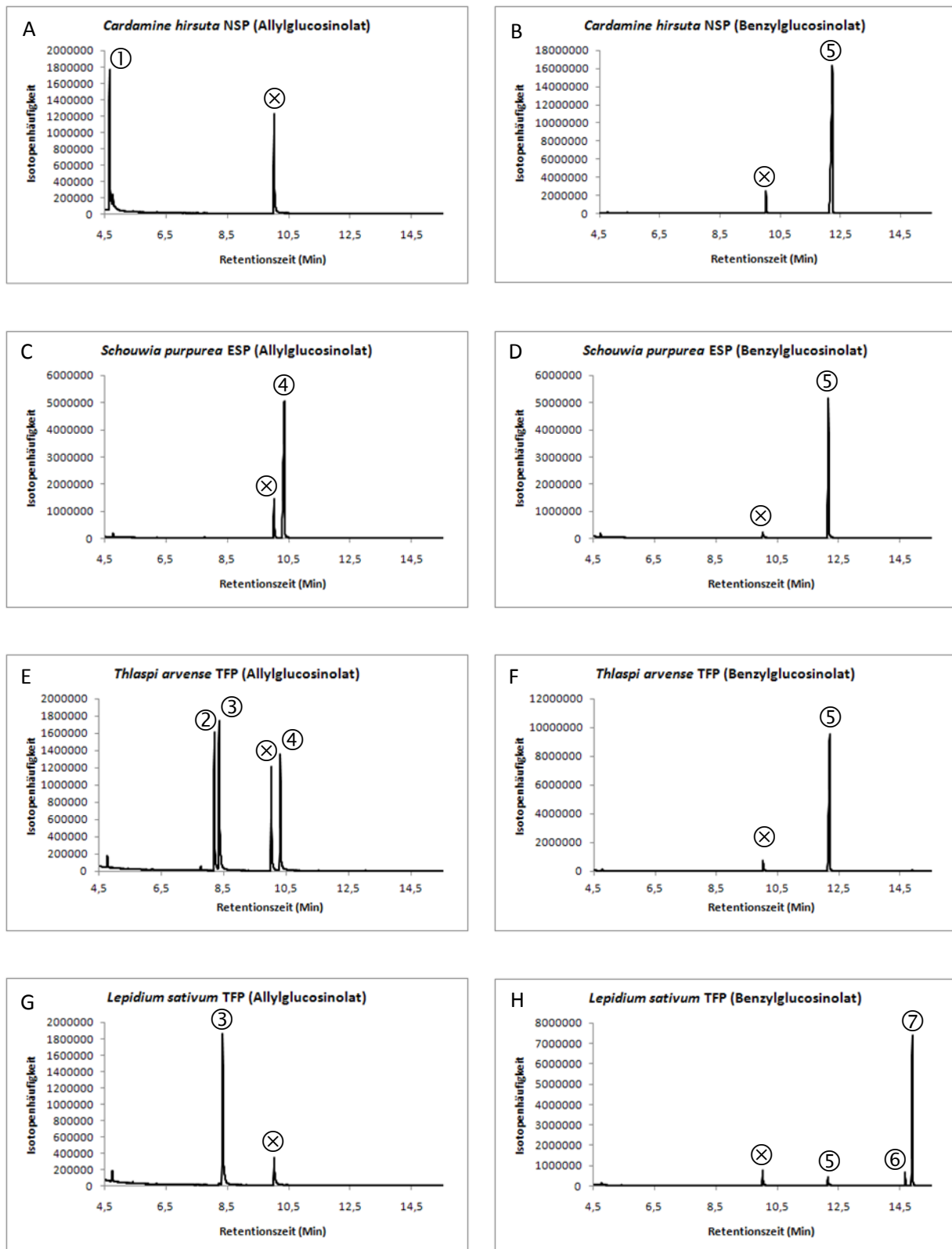


Abbildung 9: Spezifität einiger der identifizierten spezifizierenden Proteine. cDNAs identifizierter spezifizierender Proteine wurden in *E. coli* zur Expression gebracht und bakterielle Rohextrakte oder daraus aufgereinigte Proteine in Assays mit jeweils Allyl- oder Benzylglucosinolat und Myrosinase inkubiert. Erhaltene Glucosinolat-Hydrolyseprodukte wurden mit Dichlormethan extrahiert und durch GC-MS-Analytik identifiziert. Gezeigt sind repräsentative GC-MS-Chromatogramme der Dichlormethanextrakte aus Assays unter Einsatz spezifizierender Proteine der Spezies *C. hirsuta* (A, B), *S. purpurea* (C, D), *T. arvense* (E, F) und *L. sativum* (G, H) als Referenz eines zuvor bekannten TFP mit abweichender Produktspezifität. Nicht markierte Peaks entsprechen nicht relevanten Hintergrundprodukten, welche der abgelaufenen Myrosinase-Reaktion nicht zuzuordnen sind. Weitere Peaks entsprechen: ① = einfaches Nitril aus Allylglucosinolat, ② = Thiocyanat aus Allylglucosinolat, ③ = Isothiocyanat aus Allylglucosinolat, ④ = Epithionitril aus Allylglucosinolat, ⑤ = einfaches Nitril aus Benzylglucosinolat, ⑥ = Thiocyanat aus Benzylglucosinolat, ⑦ = Isothiocyanat aus Benzylglucosinolat, X = interner Standard Phenylcyanid

Tabelle 6: Spezifität verfügbarer spezifizierender Proteine. Angegeben sind die Seitenketten der Glucosinolate, die bei der Myrosinase-katalysierten Glucosinolat-Hydrolyse unter Einfluss des jeweiligen Proteins in die angegebenen Typen von Glucosinolat-Hydrolyseprodukten umgewandelt werden. TFP = Thiocyanat-formendes Protein, ESP = Epithiospezifizierendes Protein, NSP = Nitril-spezifizierendes Protein, Allyl = Allylglucosinolat, Benzyl = Benzylglucosinolat.

Spezies	Bezeichnung	Epithionitril	Nitril	Thiocyanat
<u>Im Rahmen dieser Arbeit identifizierte spezifizierende Proteine</u>				
<i>Alliaria petiolata</i>	ApTFP	Allyl	Benzyl	Allyl
<i>Cardamine hirsuta</i>	ChESP	Allyl	Benzyl	-
	ChNSP	-	Allyl / Benzyl	-
<i>Cardamine impatiens</i>	CiESP	Allyl	Benzyl	-
<i>Draba aurea</i>	DrESP	Allyl	Benzyl	-
<i>Draba lanceolata</i>	DIESP	Allyl	Benzyl	-
<i>Isatis tinctoria</i>	ItESP	Allyl	Benzyl	-
	ItNSP	-	Allyl / Benzyl	-
<i>Schouwia purpurea</i>	SpESP	Allyl	Benzyl	-
	SpNSP	-	Allyl / Benzyl	-
<i>Thlaspi arvense</i>	TaTFP	Allyl	Benzyl	Allyl
<u>Im Vorfeld bekannte spezifizierende Proteine (Literaturangaben siehe Tabelle 1)</u>				
<i>Arabidopsis thaliana</i>	AtESP	Allyl	Benzyl	-
	AtNSP1	-	Allyl / Benzyl	-
	AtNSP2	-	Allyl / Benzyl	-
	AtNSP3	-	Allyl / Benzyl	-
	AtNSP4	-	Allyl / Benzyl	-
	AtNSP5	-	Allyl / Benzyl	-
<i>Brassica oleracea</i>	BoESP	Allyl	Benzyl	-
<i>Lepidium sativum</i>	LsTFP	Allyl	Benzyl	Benzyl

2.5 Datenbankrecherche zur Identifizierung putativer spezifizierender Proteine

Um weitere Sequenzen spezifizierender Proteine aufzufinden, wurde eine Durchmusterung der Nukleotiddatenbank des NCBI Webservers mittels der blastn-Funktion (Altschul et al. 1990; Altschul et al. 1997) unter Verwendung vorhandener Nukleotidsequenzen (Tabelle 1, Tabelle 6) als Suchanfrage durchgeführt. Unter Einsatz der Nukleotidsequenz des BoESP ergab die Eingabe per Megablast-Option eine Identität von 87 % zu einer funktionell nicht charakterisierten *Full-length* Nukleotidsequenz aus *Brassica rapa* ssp. *pekinensis* (*B. rapa*; Brassicaceae) (FJ374638, Park, Y.-D. und Park, J.-H. unveröffentlicht) und eine Übereinstimmung von 78 % zu einer Sequenz aus *A. lyrata* ssp. *lyrata* (XM_002884120, Bakker E., Bergelson J., et al., unveröffentlicht). Eine Suchanfrage mit AtNSP5 identifizierte eine Nukleotidsequenz aus *Eutrema halophilum* (*E. halophilum*; Brassicaceae) (AK353179 (Taji et al. 2008)) mit 98 % Identität zu der angefragten Sequenz. Weitere Treffer wurden wegen zu geringer Übereinstimmungen oder nur unvollständig vorhandener ORFs nicht in die weiteren phylogenetischen Untersuchungen einbezogen. Zur besseren Übersicht wurden die mittels BLAST-Suche identifizierten putativen spezifizierenden Proteine in Tabelle 7 zusammengefasst.

Tabelle 7: Nukleotidsequenzen mit Ähnlichkeit zu cDNAs von ESPs und NSPs: Mittels der blastn-Funktion der Nukleotiddatenbank des NCBI Webservers an Hand vorhandener Nukleotidsequenzen als Suchanfrage identifizierte putative spezifizierende Proteine. Unter Bezeichnung ist die auf Grund der Sequenzähnlichkeit vermutete Aktivität der entsprechenden Proteine angegeben.

Spezies	Bezeichnung	Tribus	Accession-nummer	Referenz
<i>Brassica rapa</i>	ESP	Brassicaceae	FJ374638	unveröffentlicht
<i>Arabidopsis lyrata</i>	ESP	Camelineae	XM_002884120	unveröffentlicht
<i>Eutrema halophilum</i>	NSP	Eutremeae	AK353179	(Taji et al. 2008)

Somit konnte der anfängliche Sequenzdatensatz ausgehend von acht im Vorfeld bekannten spezifizierenden Proteinen um elf biochemisch charakterisierte und drei putative spezifizierende Proteine erweitert werden. Der gesamte Datensatz umfasste somit 22 cDNAs sowie die ähnliche Sequenz At3g07720 (Burow et al. 2009; Kissen and Bones 2009). Die cDNAs waren acht Tribus aus den zwei Abstammungslinien I und II der Familie Brassicaceae zuzuordnen (Al-Shehbaz et al. 2006; Beilstein et al. 2006) (Abstammungslinie I: Lepidieae, Camelineae, Cardamineae; Abstammungslinie II: Arabideae, Brassiceae, Eutremeae, Isatideae, Thlaspidiae). Hierbei ist zu vermerken, dass an diesem Punkt sowie im weiteren Verlauf der Arbeit die Abstammungslinie II zusammengefasst mit ihrer Erweiterung betrachtet wird.

2.6 Sequenzanalyse spezifizierender Proteine

2.6.1 Sequenzidentitäten

Eine Identitätsanalyse der 23 einbezogenen Sequenzen aus der Familie der Brassicaceae (Tabelle 1, Tabelle 6, Tabelle 7) mit Hilfe des EBI Webservers für ClustalW2 (Chenna, Sugawara et al. 2003) legte dar, dass die sechs im Rahmen dieser Arbeit charakterisierten ESPs und die zwei putativen ESPs der Datenbankrecherche im Vergleich mit AtESP eine hohe Sequenzidentität zwischen 80 % und 85 % auf Nukleotidebene und von 76 % bis 83 % auf Aminosäureebene zeigten (Tabelle S 2). Des Weiteren zeigten Vergleiche aller ESPs mit NSPs geringere Ähnlichkeiten von 45 % - 60 % (Aminosäuren) und 61 % - 69 % (Nukleotide) (Tabelle S 2).

Vergleiche der drei neu identifizierten NSPs lieferten in Gegenüberstellung mit AtNSP5 und At3g07720 (Burow et al. 2009; Kissen and Bones 2009) geringere Sequenzidentitäten (Nukleotide: 66 % - 67 %, Aminosäuren: 51 % - 53 %) als ein Vergleich mit AtNSP1 - AtNSP4 (Nukleotide: 82 % - 88 %, Aminosäuren: 74 % - 84 %) (Tabelle S 2). Im Gegensatz dazu hatte die der Datenbankrecherche entnommene homologe Sequenz von AtNSP5 aus *E. halophilum* (im Weiteren EhNSP genannt) eine hohe Ähnlichkeit von 84 % auf Nukleotidebene und von 85 % auf Aminosäureebene mit AtNSP5, hingegen nur eine Ähnlichkeit von 67 % - 70 % (Nukleotide) und 52 % - 54 % (Aminosäuren) mit AtNSP1 - AtNSP4 (Tabelle S 2). Die Identität dieses Homologen mit At3g07720 lag ebenfalls bei einem geringeren Wert von 65 % (Nukleotide) bzw. 57 % (Aminosäuren) (Tabelle S 2).

Eine Gegenüberstellung der zwei im Rahmen der vorliegenden Arbeit charakterisierten TFPs mit LsTFP ergab eine Identität von 73 % auf Nukleotidebene und 59 % (TaTFP) bis 60 % (ApTFP) auf Aminosäureebene (Tabelle S 2). Interessanterweise lagen die Ähnlichkeiten in einem Vergleich aller drei TFPs mit AtESP höher als im Vergleich von ApTFP und TaTFP mit LsTFP (Nukleotide: 78 % - 79 % Aminosäuren: 68 % - 72 %) (Tabelle S 2). Eine sehr hohe Sequenzidentität von 94 % auf Nukleotidebene und 93 % auf Aminosäureebene zeigten ApTFP und TaTFP (Kuchernig et al. 2011) untereinander (Tabelle S 2). Weitere Vergleiche aller drei TFPs mit AtNSP1 bis AtNSP5 und At3g07720 zeigten Sequenzidentitäten zwischen 62 % und 68 % auf Nukleotidebene und zwischen 44 % und 63 % auf Aminosäureebene (Tabelle S 2).

2.6.2 Strukturanalysen

Bezugnehmend auf Burow et al. 2009 wurden die neu identifizierten 14 Sequenzen spezifizierender Proteine (Tabelle 6, Tabelle 7) mittels der InterProScan Sequence Search Option des EMBL-EBI Webserver (Zdobnov and Apweiler 2001) auf Strukturmotive hin analysiert.

Alle im Rahmen der vorliegenden Arbeit charakterisierten und die der Datenbankrecherche entstammenden ESPs hatten laut Vorhersage eine, dem im Vorfeld bekannten AtESP (Lambrix et al. 2001) identische Struktur aus vier Kelch-Motiven (Abbildung 10; vergleiche Abbildung 4). Ebenfalls auf eine vorhergesagte Domänenstruktur bestehend aus vier Kelch-Motiven lassen sich die neu identifizierten TFPs zurückführen. Auch diese sekundären Strukturelemente stimmen mit denjenigen des bereits im Vorfeld bekannten LsTFP (Burow et al. 2007a) überein (Abbildung 10; vergleiche Abbildung 4). Für die neu charakterisierten NSPs ergab die Vorhersage vier Kelch-Motive entsprechend der bekannten *NSP*-Sequenzen (Tabelle 1) (Abbildung 10; vergleiche Abbildung 4), sowie jeweils eine N-terminal Lektin-ähnliche Jacalindomäne (Abbildung 10; vergleiche Abbildung 4). Diese gehört zum gleichen Typ wie die in AtNSP1 – AtNSP4 gefundene Lektin-ähnliche Jacalindomäne. Eine vorhergesagte Domänenstruktur entsprechend At3g07720 (Burow et al. 2009; Kissen and Bones 2009) und AtNSP5 aus drei Kelchdomänen ohne vorgelagerte Lektin-ähnliche Jacalindomäne wies die der Datenbankrecherche entstammende Sequenz EhNSP auf (Abbildung 10; vergleiche Abbildung 4).

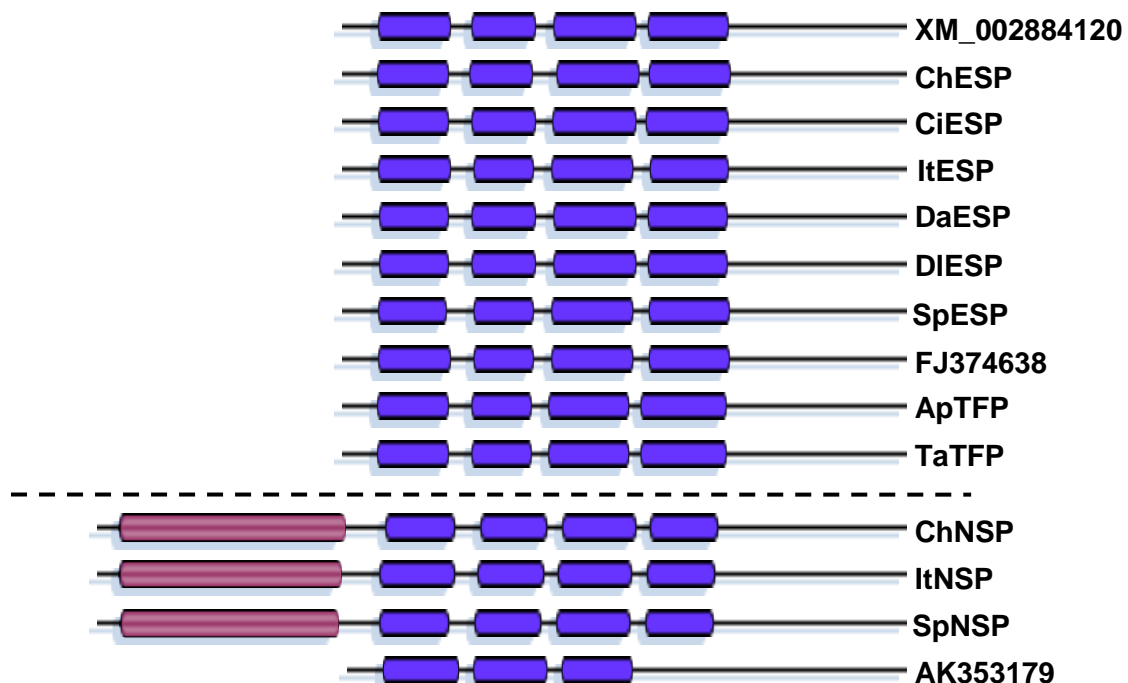


Abbildung 10: Vorhersage der Domänenstruktur neu identifizierter spezifizierender Proteine. Mittels InterProScan (EMBL-EBI Webserver) identifizierte Kelchdomänen (blaue Kästen) und Lektin-ähnliche Jacalindomänen (violette Kästen) der neu identifizierten spezifizierenden Proteine wurden dargestellt.

2.7 Phylogenetische Analyse zur Aufklärung der evolutiven Abstammung spezifizierender Proteine

2.7.1 Alignment

Die Basis einer phylogenetischen Analyse bildet ein korrektes Alignment der Sequenzdaten. Dabei werden die Sequenzdaten entsprechend einer Matrix, zum Beispiel „*Percent Accepted Mutations-Matrix*“ (PAM), „*Blocks Substitution Matrix*“ (BLOSUM) oder Gonnet, arrangiert. Für das Alignment der 23 cDNAs aus der Familie Brassicaceae wurden die Nukleotidsequenzen zunächst in Aminosäuresequenzen translatiert. Das paarweise Alignment wurde dann auf Ebene der Aminosäuren ausgeführt. Auf diese Weise wurde sichergestellt, dass der Tripletcode der Nukleotide nicht durch das Setzen von Lücken im Alignment, die nicht einem Vielfachen von drei entsprechen, verschoben wurde. Der Leserahmen der Sequenzen blieb bei der finalen Rücktranslation in Nukleotide erhalten. Zudem sind Aminosäuresequenzen verlässlicher in einem Alignment zu vergleichen, da Aminosäuren langsamer divergieren als Nukleotide. Für die phylogenetischen Analysen verwendete Alignments wurden mit den im Programm MEGA 5 (Tamura et al. 2011) implementierten ClustalW- (Thompson et al. 1994) und Muscle- (Edgar 2004) Optionen erstellt. Verwendet wurden die vorgegebenen Standardeinstellungen des jeweiligen Programmes unter Variation der zur Verfügung stehenden Matrizen. Alle auf diese Weise erhaltenen paarweisen Alignments wurden untereinander verglichen, und Lücken im Alignment wurden als am höchsten wahrscheinlich akzeptiert, wenn diese wiederholt an identischen Positionen auftraten. Um ein korrektes Alignment zu erhalten, wurden nach Abgleich der Lückenpositionen zusätzliche Korrekturen per Hand vorgenommen. Die Basis hierfür lieferte die Matrix PAM250 (Dayhoff et al. 1978).

2.7.2 Substitutionsmodelle

Über die Jahre hat sich aus dem einfachen und unkorrigierten Abzählen der zu beobachtenden Unterschiede zweier Sequenzen (p -Distanz) eine Vielzahl an Substitutionsmodellen für Distanzverfahren mit zunehmend komplexeren Annahmen zur Berechnung phylogenetischer Bäume entwickelt. Ziel dieser Verfahren ist es, berücksichtigen zu können, wie viele Substitutionen eine bestimmte Position in einer Sequenz im Verlauf der Evolution tatsächlich durchlaufen hat. Dabei liefert das genaueste Modell mit den meisten Parametern nicht zwingend die besten Ergebnisse, da durch die vielen einbezogenen Faktoren ein empirischer Fehler entstehen kann, der zu falschen Daten führt. Zu wenige Parameter können hingegen ebenfalls ungenaue Resultate zur Folge haben. Um festzustellen, welches das am besten geeignete Modell für den vorliegenden Datensatz darstellt, wurde das erstellte Alignment den im Programm MEGA 5 integrierten Möglichkeiten zum Vergleich der Substitutionsmodelle unterzogen. Für die Distanzberechnung auf Nukleotidebene wurde das Kimura-2-Parameter-Modell (Kimura 1980) vorgeschlagen. Dieses Modell geht von den zwei Substitutionswahrscheinlichkeiten (erwartete Anzahl an Substitutionen pro Sequenzposition entlang eines Zweiges bzw. über

die Zeit) der Transversion (Austausch von Purin- und Pyrimidinbasen gegeneinander) und der Transition (Austausch von Purin- oder Pyrimidinbasen untereinander) aus. Somit wird berücksichtigt, dass ein Nukleotid auf drei Wegen zu drei anderen Nukleotiden substituiert werden kann, wobei zwei Wege davon Transversionen sind und nur einer eine Transition darstellt. Das Modell geht außerdem davon aus, dass alle Nukleotide mit gleicher Frequenz auftreten (Knoop V. 1980). Dieses Modell kann mit einer Korrektur für invariable Positionen (+I) und für positionsspezifisch variable Raten (+G) ergänzt werden. Auf Ebene der Aminosäuren wurde das WAG-Modell (Whelan and Goldman 2001) von MEGA 5 angeboten. Diesem Modell liegt eine Substitutionswahrscheinlichkeits-Matrix zu Grunde, welche direkt aus den zu analysierenden Daten über eine Maximum-Likelihood-Schätzung bestimmt wird. Jedem Übergang einer Aminosäure in eine andere wird eine eigene Wahrscheinlichkeit zugeordnet. Da die Anzahl zu schätzender Parameter im Fall von Aminosäuren (189 Austauschraten und 19 relative Häufigkeiten) sehr hoch ist, funktioniert dieses Prinzip nur bei besonders großen Datensätzen zuverlässig (Knoop V. 1980). Für Protein-codierende Nukleotidsequenzen wurde von Whelan und Goldman eine solche Datenmatrix aus 3905 Aminosäuresequenzen bestimmt und als WAG-Modell benannt. Auch hier werden positionsspezifisch variable Raten (+G) einbezogen, jedoch wird in Anlehnung an die von MEGA 5 vorgeschlagene Berechnungsvariante die Korrektur der Anteile invariabler Positionen (+I) außer acht gelassen.

Auf Basis der ausgewählten Distanzmatrizes erfolgte die Rekonstruktion der Phylogenie der spezifizierenden Proteine unter Verwendung verschiedener Methoden (siehe Abschnitt 2.7.3). Als Außengruppe wurde für jede Stammbaumberechnung das Protein CAO41153 (Jaillon et al. 2007) aus *Vitis vinifera* (Vitaceae) genutzt. Die Ordnung Vitales, in die die Vitaceae eingeordnet werden, gehört wie die Ordnung Brassicales zu den Dikotyledonen, jedoch hier zu den Eurosiden I. Sie besitzt somit eine ausreichende evolutive Distanz zu den zu untersuchenden Spezies der Ordnung Brassicales. Zudem kommen Glucosinolate in den Vitales nicht vor. Die Eignung des funktionell nicht charakterisierten Proteins CAO41153 als Außengruppe erklärt sich durch seine Aminosäuresequenzidentität von 66 % mit At3g07720 (Burow et al. 2009; Kissen and Bones 2009), das trotz seiner Sequenzähnlichkeit zu NSPs keine Aktivität als spezifizierendes Protein aufweist.

2.7.3 Generation phylogenetischer Stammbäume

2.7.3.1 Betrachtungen der Nukleotidsequenzen

Bei dem, für die Konstruktion des in Abbildung 11 dargestellten Stammbaumes verwendeten Neighbor Joining-Verfahren (Saitou and Nei 1987), handelt es sich um ein Distanzverfahren, bei welchem die gegebenen alinierten Sequenzdaten in eine Distanzmatrix nach vorgegebenem Substitutionsmodell (siehe oben) konvertiert werden. Diese Methode basiert auf der *Minimum Evolution Theorie* (Camin 1965), welche besagt, dass derjenige Baum der korrekte ist, welcher in der Summe die kürzesten Astlängen liefert. Zudem ist diese Methode unabhängig von der Annahme, dass die Divergenz zwischen zwei Sequenzen proportional zur Zeit sei (molekulare Uhr; (Zuckerlandl and Pauling 1965)).

Als weitere Variante wurde die Charakter-orientierte Maximum Likelihood-Methode (Fisher 1912; Aldrich 1997) genutzt (Abbildung 12). Hier wird nach derjenigen Baumtopologie gesucht, die die höchste Wahrscheinlichkeit zeigt, den beobachteten Datensatz (Alignment) an Hand eines vorgegebenen Substitutionsmodells (siehe oben) zu reflektieren.

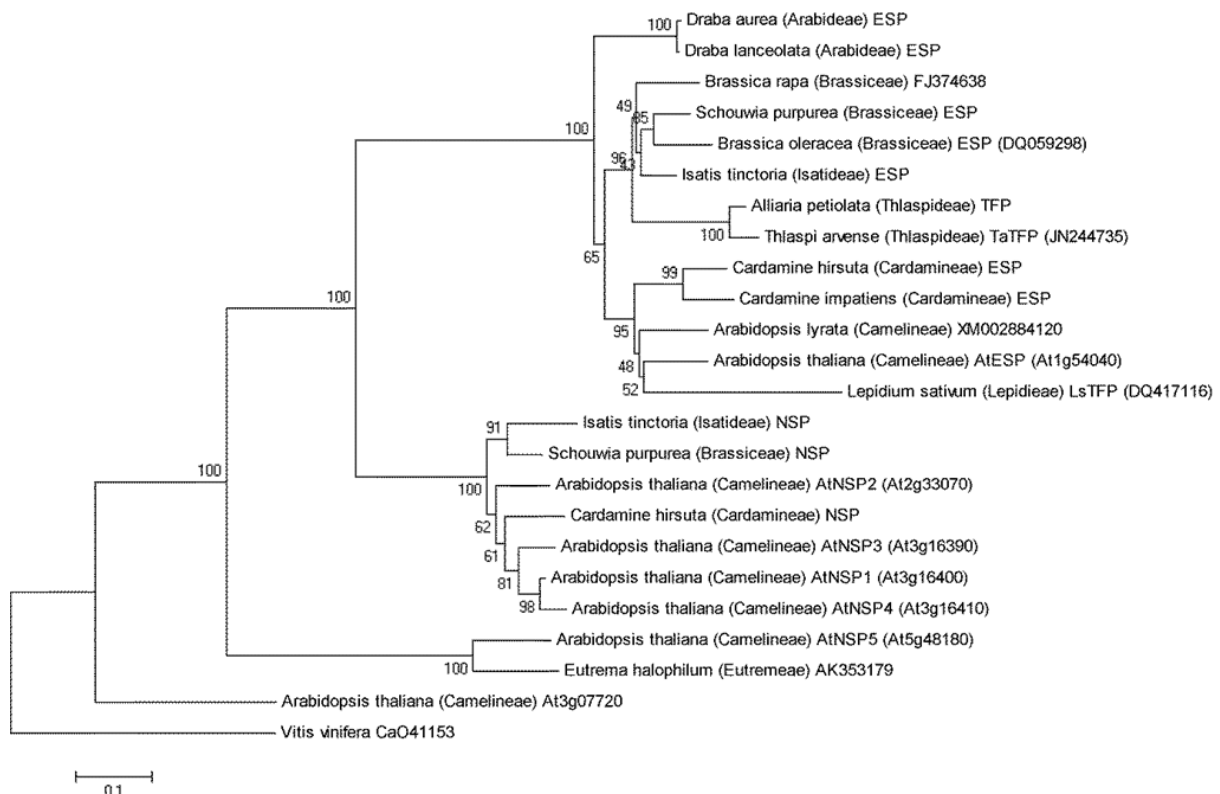


Abbildung 11: Neighbor Joining Phylogenie spezifizierender Proteine. Zur Erstellung verwendet wurden die Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der Neighbor Joining Baum basiert auf dem Kimura-2-Parameter-Modell und ist angegeben mit 1000facher Bootstrap Wiederholung. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Astlängen geben die geschätzten evolutiven Distanzen wieder. Die Maßstableiste zeigt die Länge für die angegebene Anzahl an Substitutionen pro Position.

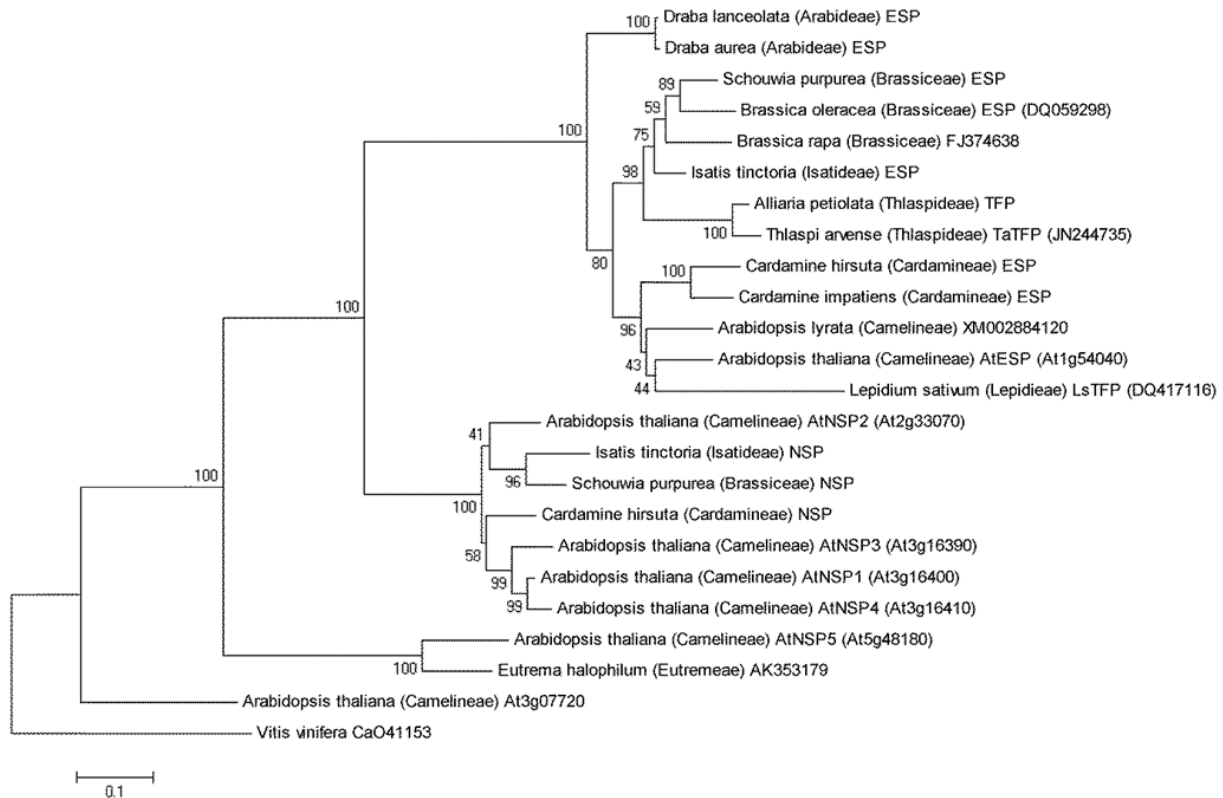


Abbildung 12: Maximum Likelihood Phylogenie spezifizierender Proteine. Zur Erstellung verwendet wurden die Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der Maximum Likelihood Baum basiert auf dem Kimura-2-Parameter-Modell und ist angegeben mit 1000facher Bootstrap Wiederholung. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Astlängen geben die geschätzten evolutiven Distanzen wieder. Die Maßstableiste zeigt die Länge für die angegebene Anzahl an Substitutionen pro Position.

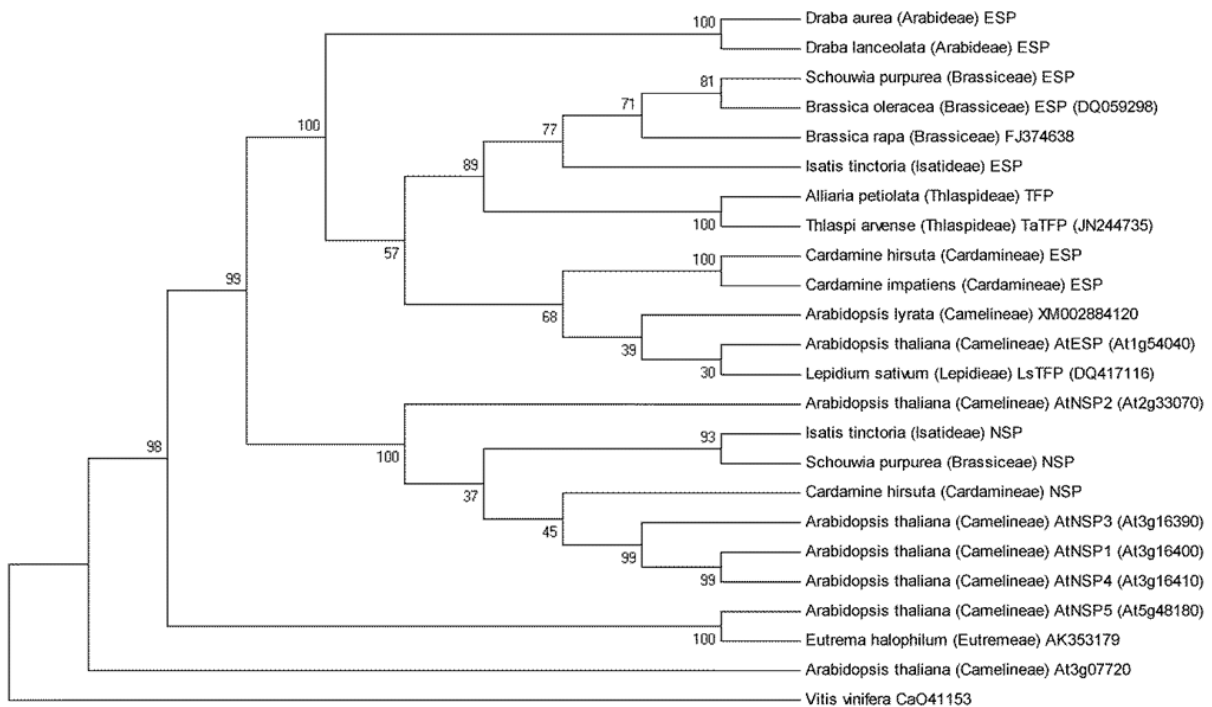


Abbildung 13: Maximum Parsimonie Phylogenie spezifizierender Proteine. Zur Erstellung verwendet wurden die Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der Maximum Parsimonie Baum basiert auf dem Kimura-2-Parameter-Modell und ist angegeben mit 1000facher Bootstrap Wiederholung. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Astlängen geben die geschätzten evolutiven Distanzen wieder.

Abbildung 13 zeigt die mit MEGA 5 (Tamura et al. 2011) erstellte Topologie eines phylogenetischen Baumes nach dem Parsimonieprinzip (William of Ockham * um 1285, † 9. April 1347). Das Parsimonieprinzip (Prinzip der minimalistischen Ökonomie) besagt, dass ein einfaches Modell einem komplexen Modell vorzuziehen ist, wenn es eine Beobachtung erklären kann. Mit der Maximum-Parsimoniemethode (Camin 1965) wird derjenige Stammbaum gesucht, welcher sich in der Summe durch die geringste Anzahl an Änderungen aller Merkmale (Merkmal = Spalte im Alignment; Merkmalszustand = Nukleotidaustausche pro Position) ausweist. Diese Methode stellt eine ebenfalls Charakter-orientierte Methode zur Berechnung von Stammbäumen dar. Sie kommt ohne Distanzmatrix aus und extrahiert die nötige Information direkt aus dem Alignment. Unter Umständen liefert die Parsimoniemethode nicht einen „optimalen“ Stammbaum sondern mehrere Bäume identischer Länge.

Die Auswahl der Stammbäume erfolgte entsprechend der besten Bootstrapwerte (Efron 1979). Diese dienen der Abschätzung des statistischen Fehlers der erstellten Baumtopologie und geben somit die Verlässlichkeit eines Knotenpunktes innerhalb eines Stammbaumes an. Die mit den drei beschriebenen Verfahren auf Basis der Nukleotidsequenzen erstellten Stammbäume zeigen eine vergleichbare Topologie mit grundlegend gleichem Clusterverhalten der Sequenzen (Abbildung 14). Diese Übereinstimmung weist auf eine hohe Zuverlässigkeit der vorhergesagten Verwandtschaftsverhältnisse der spezifizierenden Proteine hin.

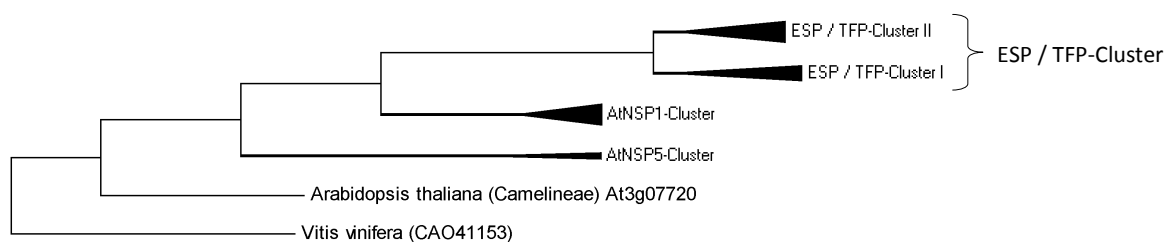


Abbildung 14: Phylogenie spezifizierender Proteine in schematischer Darstellung. Zur Erstellung verwendet wurden die Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der schematisch dargestellte Baum basiert auf den Topologien des Neighbor Joining Baumes, des Maximum Likelihood Baumes und des Maximum Parsimonie Baumes jeweils basierend auf dem Kimura 2-Parameter-Modell. CAO41153 aus *V. vinifera* diene als Außengruppe zur Bewurzelung. Die Darstellung zeigt das grundlegende Clusterverhalten der Sequenzen. Astlängen geben die geschätzten evolutionären Distanzen wieder.

Ausgehend von dem basal stehenden At3g07720 (Burow et al. 2009; Kissen and Bones 2009) zeigen sich drei Cluster spezifizierender Proteine (Abbildung 14). Das erste und abstammungsgeschichtlich älteste Cluster (AtNSP5-Cluster) bildet hierbei eine Gruppe aus zwei Sequenzen, dem putativen EhNSP (Eutremeae) (Taji et al. 2008) und dem AtNSP5 (Camelineae) (Burow et al. 2009; Kissen and Bones 2009) (Abbildung 14). Ausgehend von dieser Gruppe erscheint eine Aufspaltung in eine, ausnahmslos weitere NSPs umfassende Gruppe (AtNSP1-Cluster) und eine ausschließlich ESPs und TFPs umfassende Gruppe (ESP / TFP-Cluster) (Abbildung 14). Im AtNSP1-Cluster treten die NSPs aus *A. thaliana*, *Cardamine hirsuta* (*C. hirsuta*), *I. tinctoria* und *Schouwia purpurea* (*S. purpurea*) auf. Sie schließt somit die Tribus Camelineae, Cardamineae, Isatideae und Brassiceae ein. Auch das

ESP / TFP-Cluster enthält ESPs aus *A. thaliana*, *C. hirsuta*, *I. tinctoria* und *S. purpurea*. Zusätzlich sind ihm jedoch ESPs aus *A. lyrata*, *B. oleracea*, *B. rapa*, *Cardamine impatiens* (*C. impatiens*), *Draba aurea* (*D. aurea*), und *Draba lanceolata* (*D. lanceolata*) sowie TFPs aus *A. petiolata*, *Lepidium sativum* (*L. sativum*) und *T. arvense* zuzuordnen. Daher beinhaltet das ESP / TFP-Cluster neben den Spezies der Tribus Camelinaeae, Cardamineae, Isatideae und Brassiceae auch Spezies der Tribus Thlaspidiae und Lepidieae. Bei genauerer Betrachtung dieses ESP / TFP-Clusters lassen sich zwei Schwestergruppen erkennen, welche sich nicht in eine ESPs enthaltende Gruppe und eine TFPs enthaltende Gruppe auftrennen. Stattdessen ergibt sich eine Auflösung nach Tribus (Abbildung 11, Abbildung 12, Abbildung 13). Hierbei gruppieren jeweils spezifizierende Proteine aus Arten der Abstammungslinie I (Camelinaeae, Cardamineae, Lepidieae) und II (Arabideae, Brassiceae, Isatideae, Thlaspidiae) (Al-Shehbaz et al. 2006; Beilstein et al. 2006) zusammen. Die TFP-Sequenzen stehen sich dabei über beide Gruppen gesehen als Paraphylum gegenüber (Abbildung 11, Abbildung 12, Abbildung 13).

2.7.3.2 Betrachtungen der Aminosäuresequenzen

Auf Grund der Nutzbarkeit von Aminosäuren codierenden Nukleotidsequenzen wurden die bisher gezeigten phylogenetischen Stammbäume ausschließlich auf Ebene der Nukleotidsequenzen generiert. Da sich die Erstellung von phylogenetischen Stammbäumen auf Basis von Nukleotidsequenzen bei der Einbindung von Variablen jedoch sehr auf theoretische Berechnungen stützt (vergleiche 2.7.2), kann es im Fall der Verwendung von Nukleotidsubstitutionsmodellen zu einer Unterschätzung der tatsächlich auftretenden Änderungen kommen. Diese Vernachlässigung unsichtbarer Substitutionen entschärft sich während einer Analyse der Aminosäuresequenzen durch die Betrachtung von 20 Aminosäuren entgegen vier Nukleotiden. Daher wurde vergleichend ein Charakter-orientierter Stammbaum der Aminosäuresequenzen unter Verwendung der Maximum Likelihood-Methode (Fisher 1912; Aldrich 1997) und des Aminosäuresubstitutionsmodelles WAG (Whelan and Goldman 2001) generiert (Abbildung 15).

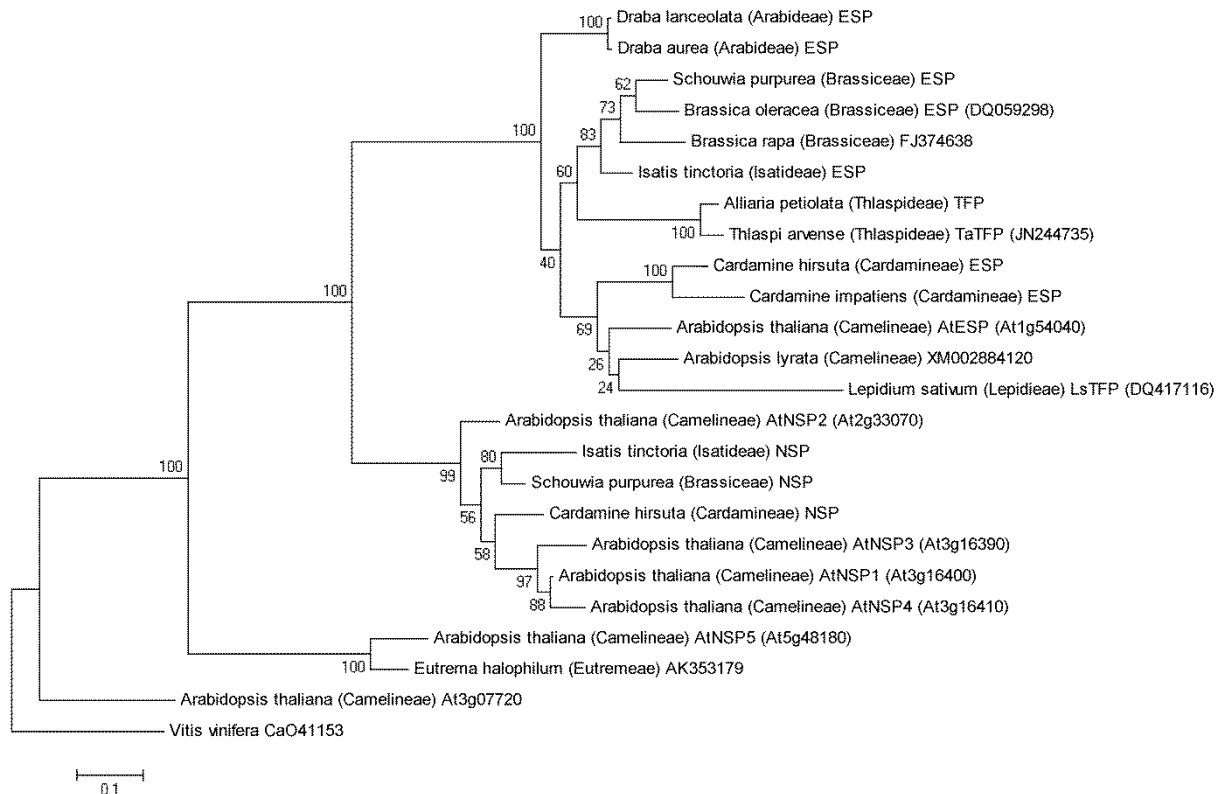


Abbildung 15: Maximum Likelihood Phylogenie spezifizierender Proteine. Zur Erstellung verwendet wurden die in Aminosäuren translatierten ORFs spezifizierender Proteine der Ordnung Brassicales. Der Maximum Likelihood Baum basiert auf dem WAG-Modell und ist angegeben mit 1000facher Bootstrap Wiederholung. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Astlängen geben die geschätzten evolutiven Distanzen wieder. Die Maßstableiste zeigt die Länge für die angegebene Anzahl an Substitutionen pro Position.

Der in Abbildung 15 dargestellte und auf Ebene der Aminosäuresequenzen berechnete Maximum Likelihood Baum entspricht in seiner grundlegenden Anordnung den Topologien der auf Ebene der Nukleotidsequenzen erstellten phylogenetischen Bäume und zeigt das beschriebene Clusterverhalten der Sequenzen (2.7.3.1).

Eine Beeinflussung dieser grundlegenden Topologie durch das zusätzliche Vorhandensein der N-terminal gelegenen Lektin-ähnlichen Jacalindomänen innerhalb des AtNSP1-Clusters (Abbildung 14) konnte theoretisch bereits durch die Tatsache ausgeschlossen werden, dass Positionen mit enthaltenen Alignmentlücken durch die gewählten Einstellungen zur Berechnung der phylogenetischen Stammbäume als uninformativ gewertet und somit ausgeschlossen wurden. Um diese Tatsache zu belegen wurde unter Verwendung des Aminosäuresubstitutionsmodelles WAG (Whelan and Goldman 2001) ein phylogenetischer Maximum Likelihood Baum unter Ausschluss der Lektin-ähnlichen Jacalindomänen berechnet (Abbildung 16). Grundlage für die nukleotidgenaue Zurückhaltung der Lektin-ähnlichen Jacalindomänen der NSPs (Tabelle 1, Tabelle 6) bildete hierbei die strukturelle Analyse der Sequenzdaten mittels der InterProScan Sequence Search Option des EMBL-EBI Webservers (Zdobnov and Apweiler 2001).

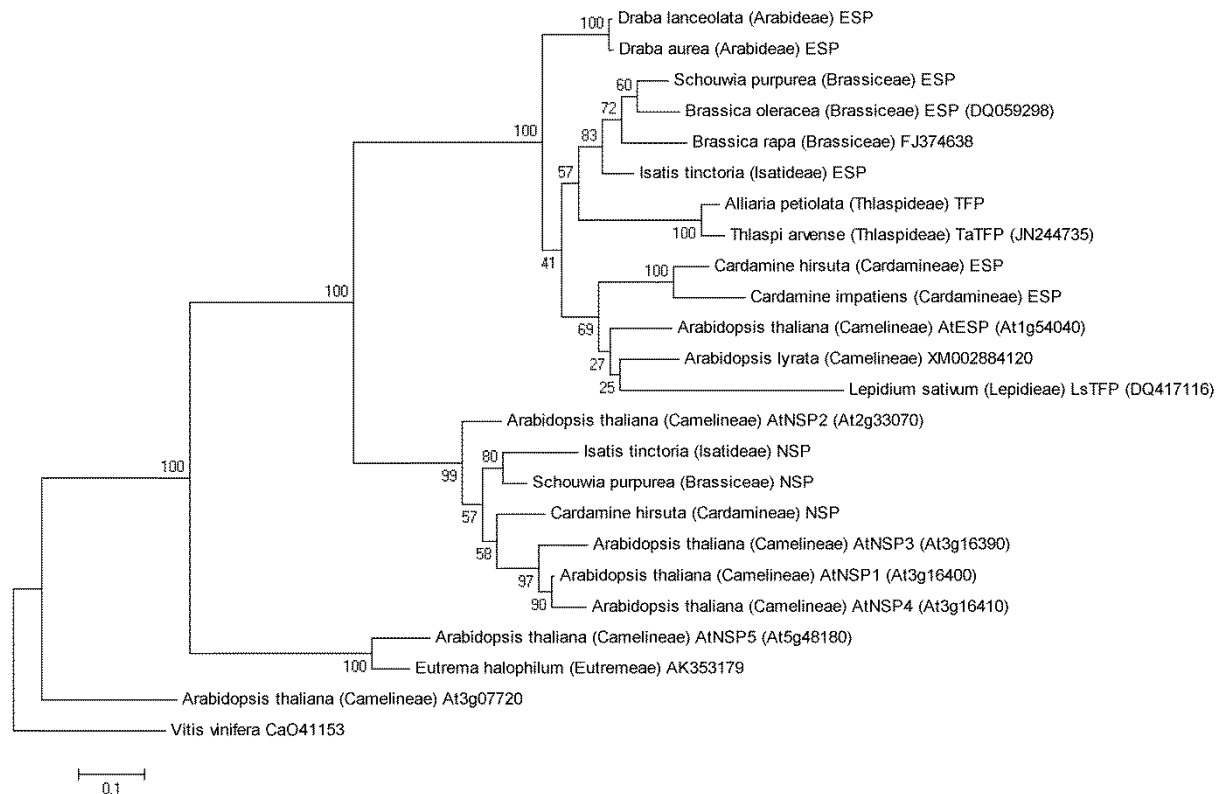


Abbildung 16: Maximum Likelihood Phylogenie spezifizierender Proteine unter nukleotidgenauem Ausschluss der Lektin-ähnlichen Jacalindomänen. Zur Erstellung verwendet wurden die in Aminosäuren translatierten ORFs von ESPs und TFPs der Ordnung Brassicales. Bei Betrachtung der NSPs wurden vorhandene Lektin-ähnliche Jacalindomänen nukleotidgenau ausgeschlossen. Der Maximum Likelihood Baum basiert auf dem WAG-Modell und ist angegeben mit 1000facher Bootstrap Wiederholung. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Astlängen geben die geschätzten evolutiven Distanzen wieder. Die Maßstableiste zeigt die Länge für die angegebene Anzahl an Substitutionen pro Position.

Die phylogenetische Untersuchung der alleinstehenden Kelchdomänen unter nukleotidgenauem Ausschluss vorhandener Lektin-ähnlicher Jacalindomänen im Fall der NSPs zeigte keine Veränderungen der grundlegenden Topologie des Stammbaumes (Abbildung 16).

Für weitere Erkenntnisse über die evolutive Entstehung der Lektin-ähnlichen Jacalindomänen und ihren Ursprung wurde in Anlehnung an Untersuchungen von Burow et al. 2009 ein phylogenetischer Baum erstellt, welcher nur die Lektin-ähnliche Jacalindomänen der NSPs betrachtet (Abbildung 17). Wie bereits in Burow et al. 2009 wurden die zwei Lektin-ähnlichen Jacalindomänen des AtNSP4 (Burow et al. 2009; Kissen and Bones 2009) getrennt voneinander einbezogen. Die N-terminal gelegene Domäne wurde dabei mit AtNSP4A, die darauf folgende mit AtNSP4B bezeichnet.

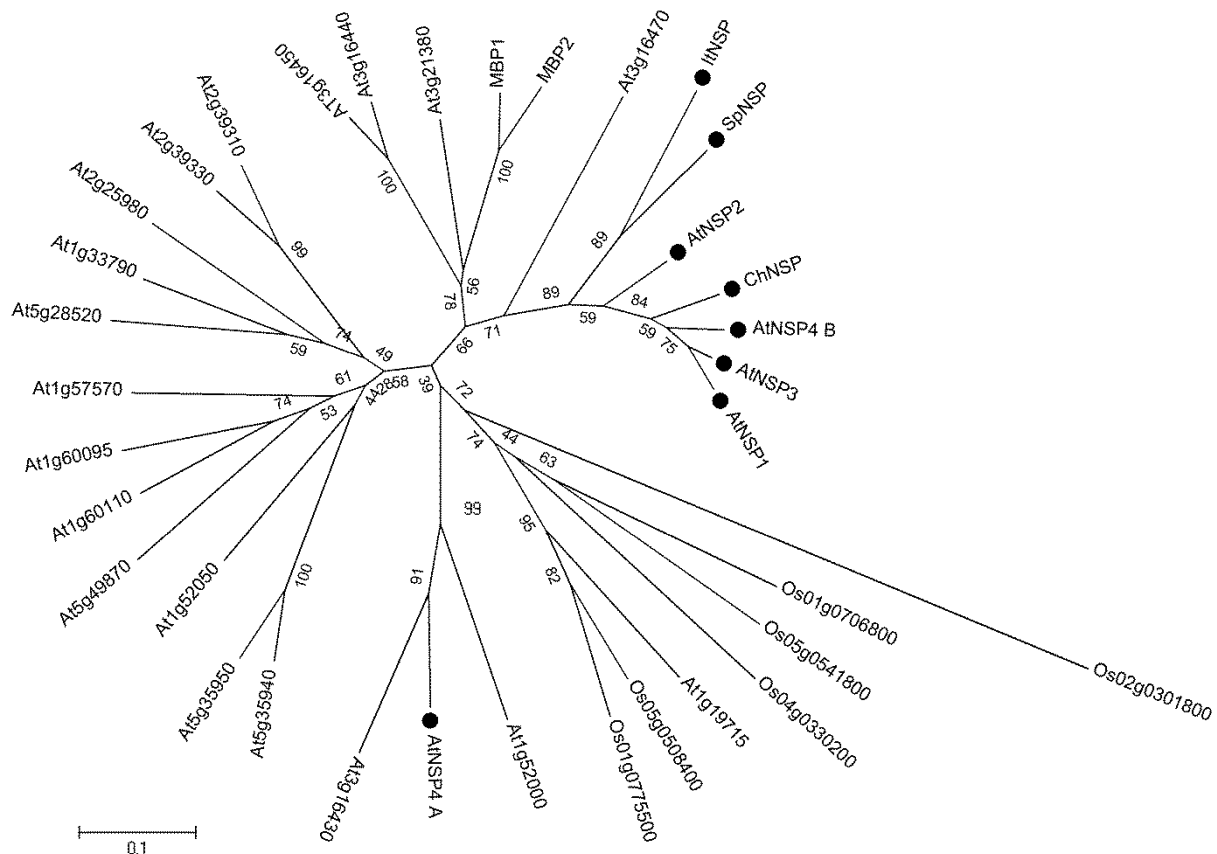


Abbildung 17: Neighbor Joining Phylogenie Lektin-ähnlicher Jacalindomänen spezifizierender Proteine. Zur Erstellung verwendet wurden die in Aminosäuren translatierten ORFs aller Proteine aus *A. thaliana* (At) und *Oryza sativa* (Os) die Lektin-ähnliche Jacalindomänen mit Ähnlichkeit zu der des MBP1 (At1g52040) aus *A. thaliana* besitzen. Für NSPs wurden ausschließlich die Lektin-ähnlichen Jacalindomänen betrachtet. Um die Herkunft einzelner Lektin-ähnlicher Jacalindomänen zu ermitteln wurden die Domänen des AtNSP4 getrennt voneinander betrachtet. Dabei stellt sich die N-terminal gelegene Domäne als AtNSP4A und die folgende Domäne als AtNSP4B dar. Der Neighbor Joining Baum ist angegeben mit 1000facher Bootstrap Wiederholung und wurde ohne Bewurzelung erzeugt um die Bestimmung der internen Verwandtschaft zu ermöglichen und die unbekannte Identität eventueller Vorgänger der Lektin-ähnlichen Jacalindomänen zu verdeutlichen. Astlängen geben die geschätzten evolutiven Distanzen wieder. Die Maßstabsleiste zeigt die Länge für die angegebene Anzahl an Substitutionen pro Position. Im Rahmen der vorliegenden Arbeit zu untersuchende Lektin-ähnliche Jacalindomänen sind durch schwarze Kreise markiert.

Die im Vorfeld dieser Arbeit bekannten Lektin-ähnlichen Jacalindomänen aus AtNSP1 - AtNSP3 sowie AtNSP4B (Burow et al. 2009) clustern wie bereits in Burow et al. 2009 beschrieben in einem Zweig mit MBP1 und MBP2 (Theologis et al. 2000). Ebenfalls in diesem Zweig gruppieren sich die Lektin-ähnlichen Jacalindomänen der neu identifizierten NSPs der Spezies *C. hirsuta*, *I. tinctoria* und *S. purpurea*. Entsprechend schließt dieser Zweig Domänen aller charakterisierten NSPs und somit die Tribus Camelineae, Cardamineae, Isatideae und Brassiceae ein. Allein die zusätzliche Lektin-ähnliche Jacalindomänen AtNSP4A des AtNSP4 fügt sich mit Abstand in einen weiteren Zweig des dargestellten Neighbor Joining Baumes funktionell nicht charakterisierter Sequenzen ein.

2.8 Paarweise Berechnung von Aminosäuredistanzen und Nukleotiddistanzen spezifizierender Proteine

Die evolutiven Änderungen zwischen zwei Sequenzen können neben der Darstellung als Stammbaum auch als Datenmatrix abgebildet werden. Eine solche Matrix ermöglicht es an Hand von berechneten Zahlenwerten den Verwandtschaftsgrad zweier Sequenzen abzuschätzen. Hierbei wird die Aminosäuredistanz (d_{aa}) definiert als die Anzahl an Aminosäuresubstitutionen über die Länge des gesamten Alignments. Eine Nukleotiddistanz (d_{na}) definiert sich entsprechend als die Anzahl an Nukleotidsubstitutionen über die Gesamtlänge des Alignments. Je weiter zwei Taxa evolutiv voneinander entfernt liegen, desto größer ist die Anzahl sie unterscheidender Nukleotid- / Aminosäureaustausche. Wie bei der Erstellung phylogenetischer Stammbäume handelt es sich bei einer derartigen Berechnung der Distanzdaten jedoch nicht um das reine Abzählen an unterschiedlichen Sequenzpositionen. Zu Grunde gelegt wurde diesen Berechnungen das oben beschriebenen Kimura-2-Parameter-Modell (Kimura 1980), ein Substitutionsmodell für Nukleotide. Im Fall der Aminosäuredistanz stellt das Programm Mega 5 (Tamura et al. 2011) nicht die bisher verwendeten Aminosäure-Substitutionsmodelle zur Verfügung sondern nimmt die Berechnung auf Basis sogenannter Codon-Substitutionsmodelle für Aminosäuren vor. Diese funktionieren ähnlich den Nukleotidsubstitutionsmodellen, nehmen jedoch zusätzlich an, dass ein Wechsel zwischen zwei Aminosäuren unwahrscheinlicher wird, je chemisch unterschiedlicher sie sich verhalten. Entsprechende Größenordnungen der Abweichung werden empirisch geschätzt. Das den vorliegenden Datensatz am besten beschreibende Modell war das Equal-input Modell. Dieses Modell nimmt eine Gleichwertigkeit der Substitutionsmuster zwischen den Verzweigungen in einem Stammbaum an. Zudem geht es zwar von einer Gleichheit der Substitutionsraten pro Sequenzposition aus, wurde jedoch für positionsspezifisch variable Raten (+G) ergänzt.

Die untersuchten NSPs (Tabelle 1, Tabelle 6, Tabelle 7) wiesen bei einem Vergleich der d_{aa} eine große Spannweite der Sequenzdivergenzen auf (5,7 % - 110,4 %) (Tabelle 8, Tabelle S 3). Hierbei zeigten At3g07720 (Burow et al. 2009; Kissen and Bones 2009), AtNSP5 und das putative EhNSP untereinander verglichen (82,7 % - 92,8 %) und verglichen mit den NSPs des AtNSP1-Clusters (AtNSP1 - AtNSP4, SpNSP, ChNSP, ItNSP) (95,5 % - 110,4 %) die größten Distanzen (Tabelle S 3). Eine Ausnahme bildete das putative EhNSP verglichen mit AtNSP5 (14,9 %) (Tabelle S 3). Alle weiteren Berechnungen innerhalb der NSPs des AtNSP1-Clusters hoben sich deutlich mit kleineren Distanzwerten von 5,7 % bis 28,6 % ab (Tabelle S 3).

Ein Vergleich aller NSPs mit ESPs zeigte Werte zwischen 60,5 % und 135,9 % (Tabelle 8, Tabelle S 3). Hierbei ergaben At3g07720, AtNSP5 und das putative EhNSP im Vergleich mit den ESPs die größten Distanzen von 107,3 % bis 135,9 % (Tabelle S 3). Die NSPs des AtNSP1-Clusters wiesen verglichen mit ESPs kleinere Distanzwerte von 60,5 % - 98,2 % auf (Tabelle S 3).

Eine Gegenüberstellung aller ESPs untereinander erbrachte Sequenzdivergenzen im Rahmen von 1,0 % (DaESP / DIESP) bis zu 41,4 % (BoESP / CiESP) (Tabelle 8, Tabelle S 3). Hier lagen die kleinsten Distanzen jeweils in Vergleichen von ESPs aus Spezies eines Tribus untereinander (Arabideae: DaESP / DIESP 1,0 %, Brassiceae: BoESP / BrESP / SpESP 11,1 % - 18,5 %, Camelineae: AIESP / AtESP 20 %, Cardamineae: ChESP / CiESP 16,2 (Tabelle S 3). ESP codierende Sequenzen verglichen mit TFP codierenden Sequenzen zeigten keine weiteren Auffälligkeiten in ihrer Spannbreite von 28,2 % bis 60,5 % (Tabelle 8, Tabelle S 3).

TFPs untereinander verglichen ergaben, bei einer Betrachtung innerhalb des Tribus Thlaspidaceae (ApTFP / TaTFP (Kuchernig et al. 2011)), eine sehr geringe Distanz von 6,5 % (Tabelle 3, Tabelle S 3). Eine Gegenüberstellung der TFPs der Tribus Lepidieae (LsTFP (Burow et al. 2007a)) und Thlaspidaceae zeigte größere Distanzen mit Werten von 68,1 % (LsTFP / ApTFP) und 70,1% (LsTFP / TaTFP) (Tabelle 8, Tabelle S 3). Diese zwei Werte liegen zudem höher als Werte aus den Vergleichen der ESPs untereinander (1,0 % - 41,4 %) (Tabelle 8, Tabelle S 3). Die Werte der TFPs in Gegenüberstellung zu NSPs wiesen eine, gegenüber dem ESP / NSP Vergleich (60,5 % - 135,9 %), leicht erhöhte Sequenzdivergenz von 79,0 % bis 150,8 % auf (Tabelle 8, Tabelle S 3). Die größten Divergenzen traten hierbei im Vergleich der *LsTFP*-Sequenz mit allen NSP codierenden Sequenzen (98,5 % - 150,8 %) und im Vergleich der Sequenzen des AtNSP5-Clusters und At3g07720 mit TaTFP und ApTFP (131,6 % - 146,1 %) auf (Tabelle S 3).

Die ebenfalls kalkulierten d_{na} Divergenzen spiegeln die hier beschriebenen Wertabweichungen in ihrer Verteilung wieder. Für eine genauere Betrachtung sind die erstellten Divergenzmatrizes der Tabelle S 3 im Anhang zu entnehmen.

Tabelle 8: Aminosäuredistanz spezifizierender Proteine. Tabellarisch aufgestellt wurden wesentliche paarweise verglichene Aminosäuredistanzen (d_{aa}) definiert als die Anzahl an Aminosäuresubstitutionen über die Länge des gesamten Alignments. Berechnungsgrundlage war das Equal-input Modell ergänzt durch positionsspezifisch variable Raten (+G). Die vollständigen Daten sind Tabelle S 3 im Anhang zu entnehmen.

Protein	TFP	ESP	NSP
TFP	6,5 %, 68,1 % + 70,1%	-	-
ESP	28,2 % - 60,5 %	1,0 % - 41,4 %	-
NSP	79,0 % - 150,8 %	60,5 % - 135,9 %	5,7 % - 110,4 %

Eine anschließende Analyse der alleinstehenden Kelchdomänen unter nukleotidgenauem Ausschluss der Lektin-ähnlichen Jacalindomänen der NSPs zeigte bei Vergleichen der ESPs und TFPs untereinander oder gegen NSPs keinerlei Veränderungen gegenüber den bisherigen Daten für die Untersuchungen der vollständigen ORFs. Diese Gegebenheit erklärt sich in der Tatsache, dass eine Position in Gegenüberstellung zu einer Alignmentlücke in der gewählten Berechnungsgrundlage als nicht informativ gewertet wird. Sich ergebende Abweichungen von d_{aa} und d_{na} zeigen lediglich leichte auf- beziehungsweise absteigende Tendenzen im Vergleich der NSPs untereinander, ergeben jedoch keine neuen Auffälligkeiten. Die entsprechenden Daten sind Tabelle S 4 zu entnehmen.

Eine Untersuchung der alleinstehenden Lektin-ähnlichen Jacalindomänen der NSPs hingegen ergab deutlich höhere Sequenzdivergenzen von 11,9 % - 249,7 % auf Ebene der Aminosäuren (d_{aa}) (Tabelle S 5) als die bisherigen Untersuchungen der vollständigen ORFs der NSPs zeigten (5,7 % - 110,4 %) (Tabelle 8, Tabelle S 3). Für die Berechnung wurden die zwei Lektin-ähnlichen Jacalindomänen des AtNSP4 getrennt behandelt. AtNSP4A bezeichnet die N-terminal gelegene Domäne und AtNSP4B die darauf folgende. Die geringste Distanz trat zwischen der Lektin-ähnlichen Jacalindomäne von AtNSP1 und AtNSP4B (11,9 %) sowie AtNSP1 und AtNSP3 (12,8 %) auf (Tabelle S 5). Die größten Distanzwerte ergeben sich aus dem Vergleich der Lektin-ähnlichen Jacalindomäne AtNSP4A mit allen weiteren Lektin-ähnlichen Jacalindomänen (164,0 % - 249,7 %) (Tabelle S 5). Das Datenspektrum aller bisher nicht aufgeführten Berechnungen konzentrierte sich im Bereich von 20,3 % - 71,0 % (Tabelle S 5).

Die auf Basis der ORFs erhobenen Daten aus den Berechnungen von d_{aa} und d_{na} spiegeln die Clusterung der dargestellten Baumtopologien (Abbildung 15) sowie die darin ersichtlichen Verwandtschaftsverhältnisse wieder. Auch die unter nukleotidgenauem Ausschluss der Lektin-ähnlichen Jacalindomänen der NSPs erhaltenen und als unverändert zu betrachtenden Werte für d_{aa} und d_{na} stehen in Übereinstimmung mit der generierten Maximum Likelihood Phylogenie unter ebenfalls nukleotidgenauem Ausschluss dieser Lektin-ähnlichen Jacalindomänen und ihrer gleichbleibenden Topologie (Abbildung 16). Des Weiteren entsprechen die berechneten Distanzwerte der alleinstehenden Lektin-ähnlichen Jacalindomänen der in Abbildung 17 unter Verwendung derselben Sequenzabschnitte erhaltenen Baumtopologie und der darin erkennbaren Einzelstellung von AtNSP4A.

2.9 Berechnung der Anzahl synonymer und nichtsynonymer Substitutionen zur Bestimmung des Selektionsdruckes

Die Berechnung der Anzahl von synonymen Substitutionen pro synonyme Sequenzposition (d_S) und der Anzahl von nichtsynonymen Substitutionen pro nichtsynonymer Sequenzposition (d_N) sowie die Berechnung des daraus folgenden Quotienten d_N / d_S offenbart das Substitutionsratenverhältnis ω ($\omega = d_N / d_S$). Dieses dient als Maß für die Art der wirkenden Selektion auf Ebene der Aminosäuren (Yang et al. 2000). Diese Tatsache erklärt sich daher, dass es sich bei synonymen Substitutionen um stille Substitutionen handelt die keinen Einfluss auf die Aminosäuresequenz haben. Daher unterliegen stille Substitutionen keiner Selektion. Hingegen sind nichtsynonyme Substitutionen der Selektion ausgesetzt da sie aminosäureaustauschend sind und zu negativen, positiven und neutralen Veränderungen der Aminosäuresequenz und damit des Phänotyps führen können. Auf nichtsynonyme Substitutionen wirkt in Fällen einer negativen Beeinflussung des Phänotyps die sogenannte negative oder auch stabilisierende Selektion die die Bereinigung sich nachteilig auswirkender Substitutionen bewirkt um die Funktionalität eines Proteins aufrecht zu erhalten. In einem solchen Fall treten daher gegenüber synonymen Substitutionen nur wenige bis keine nichtsynonymen Substitutionen auf und der berechnete Quotient ω liegt < 1 . Erweisen sich nichtsynonyme Substitutionen jedoch als vorteilhaft für den Phänotyp, so haben sie gegenüber synonymen Substitutionen eine höhere Wahrscheinlichkeit fixiert zu werden. Bei der Bewahrung vorteilhafter Substitutionen wirkt positive Selektion, welche somit auch zu funktionellen Spezialisierungen führen kann. Positive Selektion zeigt sich daher durch das Auftreten einer erhöhten Anzahl nichtsynonymer Substitutionen gegenüber synonymen Substitutionen und führt zu $\omega > 1$. Bei einem Wert von $\omega = 1$ herrscht demnach neutrale Selektion.

Auch bei der Berechnung der Parameter d_N und d_S handelt es sich nicht um das reine Abzählen ersichtlicher Substitutionen pro Position. In der vorliegenden Berechnung wurde die in dem Programm MEGA5 (Tamura et al. 2011) implementierte Möglichkeit zur Berechnung von d_N und d_S auf Basis des Kimura-2-Parameter-Modells genutzt (Kimura 1980). In den Vergleichen aller 22 spezifizierenden Proteine (Tabelle 1, Tabelle 6, Tabelle 7) und der homologen Sequenz At3g07720 (Burow et al. 2009; Kissen and Bones 2009) wurden ausschließlich ω -Werte zwischen 0,123 und 0,627 erhalten (Tabelle 9A und 9B, Tabelle S 6). Diese sprechen für negative Selektion und somit für eine Bereinigung nichtsynonymer Substitutionen innerhalb der Familie der spezifizierenden Proteine.

Negative Selektion konnte vor allem bei Gegenüberstellungen von Sequenzen für Proteine gleicher Funktionalität nachgewiesen werden. So zeigte ein Vergleich des BoESP sowohl mit dem nah verwandten ItESP ($\omega = 0,143$) als auch mit dem ebenfalls nah verwandten SpESP ($\omega = 0,198$) kleine ω -Werte (Tabelle S 6). Auch der Vergleich der TFPs aus den nah verwandten Spezies *A. petiolata* und *T. arvense* zeigte einen kleinen ω -Wert von 0,198 (Tabelle 9B, Tabelle S 6). Des Weiteren zeigte die Gegenüberstellung des EhNSP mit AtNSP5 einen ω -Wert von 0,123 (Tabelle S 6).

Ein Vergleich aller weiteren ESPs ergab ω -Werte zwischen 0,213 bis 0,627 (Tabelle S 6). Die Betrachtungen der übrigen NSPs zeigten ω -Werte von 0,239 bis 0,435 (Tabelle S 6). NSPs verglichen mit ESPs zeigten ω -Werte von 0,219 bis 0,351 (Tabelle 9A, Tabelle S 6) und eine Gegenüberstellung der NSPs mit TFPs ergab ω -Werte im gleichen Bereich von 0,234 und 0,346 (Tabelle 9A, Tabelle S 6). TFPs im Vergleich mit ESPs führten zu ω -Werten von 0,344 bis 0,549 (Tabelle 9A, Tabelle S 6). Vergleiche der drei TFPs untereinander zeigten ω -Werte von 0,198 bis 0,524 (Tabelle 9B, Tabelle S 6). Es sei vermerkt, dass hierbei der Vergleich innerhalb der TFPs aus den Spezies des Tribus Thlaspidiae (ApTFP / TaTFP) einen deutlich kleineren ω -Wert von 0,198 aufwies, während die Gegenüberstellungen des LsTFP aus einer Spezies des Tribus Lepidieae verglichen mit TaTFP ($\omega = 0,524$) sowie ApTFP ($\omega = 0,516$), jeweils aus einer Spezies des Tribus Thlaspidiae, größere ω -Werte aufwiesen (Tabelle 9B, Tabelle S 6). Für eine genauere Betrachtung sind weitere Substitutionsratenverhältnisse der Matrix in Tabelle S 6 im Anhang zu entnehmen.

Tabelle 9: Substitutionsratenverhältnisse (ω) spezifizierender Proteine. Tabellarisch dargestellt ist eine Auswahl ermittelter Substitutionsratenverhältnisse ω definiert als $\omega = d_N / d_S$ mit d_N als Anzahl nichtsynonymer Substitutionen pro nichtsynonymer Sequenzposition und d_S als Anzahl synonyme Substitutionen pro synonyme Sequenzposition. Berechnungen wurden vorgenommen auf Nukleotidsäureebene unter Verwendung des Kimura-2-Parameter-Modells. A) Paarweiser Vergleich funktionell charakterisierter TFPs. B) Paarweise Vergleiche der ESPs, TFPs und NSPs. Die vollständigen Daten sind Tabelle S 6 im Anhang zu entnehmen.

A)

Protein	TFP	ESP	NSP
ESP	0,344 – 0,549	0,143 - 0,627	
NSP	0,234 - 0,346	0,219 - 0,351	0,123 - 0,435

B)

Spezies	<i>L. sativum</i> LsTFP	<i>A. petiolata</i> TFP	<i>T. arvense</i> TaTFP
<i>Lepidium sativum</i> (Lepidieae) LsTFP	-	-	-
<i>Alliaria petiolata</i> (Thlaspidiae) TFP	0,516	-	-
<i>Thlaspi arvense</i> (Thlaspidiae) TaTFP	0,524	0,198	-

Um die Schlussfolgerung der negativen Selektion innerhalb der Familie der spezifizierenden Proteine zu untermauern, wurde ein codonbasierter Test auf positive Selektion zwischen zwei Sequenzen, der in MEGA5 als Z-Test auf Selektion bezeichnet ist, angeschlossen. Dieser Test basiert auf einem statistischen Hypothesentest, welcher das Substitutionsratenverhältnis synonyme (d_S) gegen nichtsynonyme (d_N) Substitutionen zweier Sequenzen vergleicht. Hierbei wird eine Nullhypothese (H_0) zu Grunde gelegte die von $d_S = d_N$ ausgeht, also von der Annahme, dass die Anzahl synonyme und nichtsynonyme Substitutionen identisch ist. Im Fall der Betrachtungsweise als Test auf positive Selektion lautet die Gegenhypothese (H_A) als $H_A = d_N > d_S$, da bei einer Anhäufung nichtsynonymer Substitutionen innerhalb einer Sequenz nicht mehr von einer zufälligen Verteilung, sondern

von positiver Selektion, also von einer Fixierung nichtsynonymer Substitutionen auszugehen ist. Angenommen wird im Rahmen des Z-Testes, dass es mit einer Wahrscheinlichkeit von 5 % Zufall sein könnte, dass das erhaltene Ergebnis nicht der Nullhypothese entspricht ($p = 0,05$). Liegt $p < 0,05$ so kann die Nullhypothese als signifikant abgelehnt betrachtet werden. Im Falle der untersuchten spezifizierenden Proteine ergab die Berechnung für positive Selektion betrachtet kein signifikantes Ergebnis. Die Nullhypothese konnte angenommen und somit positive Selektion abgelehnt werden (Daten nicht gezeigt). Eine Umkehr der Gegenhypothese als $H_A = d_N < d_S$, somit ein Test auf negative Selektion ergab, dass die Nullhypothese mit hoher Signifikanz abgelehnt werden konnte (Daten nicht gezeigt). Der Z-Test auf Selektion bestätigte somit die Annahme der negativen Selektion. Die einzige Ausnahme bildete die Berechnung von DaESP mit DIESP (Tabelle S 6). In diesem Fall war es nicht möglich bei $H_A = d_N < d_S$ die Nullhypothese signifikant abzulehnen (Daten nicht gezeigt). Auch ein Test auf neutrale Selektion ($H_A: d_N \neq d_S$) ließ auf Grund einer identischen Ergebnislage keine weitere Einordnung des Vergleiches zu (Daten nicht gezeigt). Auf Grund der zuvor berechneten ω -Werte < 1 für diesen Vergleich (DaESP / DIESP $\omega = 0,627$) ist jedoch ebenfalls von negativer Selektion auszugehen. Des Weiteren sind die Sequenzen sehr identisch. Das Phänomen hoher Zahlenwerte für ω kann sich in einem solchen Fall aus der geringen Anzahl erwarteter Substitutionen ergeben. Bei einer kleinen Anzahl an Aminosäureaustauschen über die gesamte Sequenz kann die Anzahl auftretender nichtsynonymer Substitutionen entgegen synonymen Substitutionen den berechneten Wert für ω trotz hoher Konservierung der Sequenz sehr hoch erscheinen lassen.

Im Folgenden wurden die Substitutionsratenverhältnisse (ω) für die alleinstehenden Kelchdomänen der NSPs unter nukleotidgenauem Ausschluss ihrer Lektin-ähnlichen Jacalindomänen berechnet. Auch hierbei ergaben sich Werte für $\omega < 1$ (0,200 bis 0,475) (Tabelle S 7). Diese entsprachen den bereits zuvor beobachteten ω -Werten für die vollständigen ORFs der NSP-Sequenzen (0,255 bis 0,435) (Tabelle 9, Tabelle S 6). Jedoch waren unter Verwendung des Z-Tests auf Selektion weder für positive und negative Selektion noch für neutrale Selektion auswertbare Ergebnisse zu erhalten (Daten nicht gezeigt).

Wurden ausschließlich die Sequenzabschnitte für die Berechnung von ω verwendet die für die Lektin-ähnlichen Jacalindomänen der NSPs codieren, konnten für die N-terminal gelegene Lektin-ähnliche Jacalindomäne des AtNSP4 (AtNSP4A) keine auswertbaren Daten erhalten werden. Eine Veränderung der Parameter würde der Vergleichbarkeit der Datensätze entgegenstehen, daher wurde die Lektin-ähnliche Jacalindomäne AtNSP4A aus den Berechnungen ausgeschlossen. Vergleiche der übrigen Lektin-ähnlichen Jacalindomänen der NSPs zeigten ω -Werte im Bereich von 0,247 bis 0,802 (Tabelle S 8). Diese Werte lagen somit deutlich höher als die zuvor für die vollständigen ORFs der NSP-Sequenzen (0,255 bis 0,435) (Tabelle 9, Tabelle S 6) berechneten Werte, wiesen jedoch mit $\omega < 1$ ebenfalls auf negative Selektion hin. Bestätigt werden konnte diese Aussage durch einen Z-Tests auf positive Selektion, bei welchem die Nullhypothese angenommen und somit positive Selektion abgelehnt werden konnte (Daten nicht gezeigt). Der Z-Test auf Selektion bestätigte somit die Annahme der negativen Selektion.

2.10 Identifizierung spezifizierender Proteine in weiteren Spezies der Ordnung Brassicales

Alle bisherigen Analysen basierten auf einem anfänglich durchgeführten phytochemischen Screening (2.1). Da in vielen Fällen bekannt ist, dass sowohl die Glucosinolatprofile als auch die Art der bei Gewebeverletzung freigesetzten Glucosinolat-Hydrolyseprodukte von Pflanzenorgan, Entwicklungszustand, Wachstumsbedingungen und Befall mit Herbivoren und Pathogenen abhängen (Brown et al. 2003; Charron et al. 2005; Charron et al. 2005; Himanen et al. 2008; Wentzell and Kliebenstein 2008; Textor and Gershenzon 2009), lieferte das phytochemische Screening (2.1, Tabelle S 1) nur eine „Momentaufnahme“. Die Möglichkeit, dass die untersuchten Spezies unter anderen Bedingungen andere Ergebnisse liefern würden und in interessanten Arten Pseudogene vorhanden sein könnten sowie die Tatsache, dass sich die Analyse bei einigen Spezies als schwierig erwies, bildeten die Grundlage der Überlegung auf genetischer Ebene auch Spezies zu untersuchen, welche im Rahmen des phytochemischen Screenings keine von Isothiocyanaten abweichenden Glucosinolat-Hydrolyseprodukte bildeten.

2.10.1 Anwendung einer PCR-Strategie

An Hand einer zum Zeitpunkt der Laborarbeiten vorliegenden Veröffentlichung (Al-Shehbaz et al. 2006; Beilstein et al. 2006) wurde eine Auswahl evolutiv aussagekräftiger Spezies getroffen. Die Wahl fiel auf die Isolation von DNA-Fragmenten aus *Cleome spinosa* (*C. spinosa*), *Limnanthes douglasii* (*L. douglasii*) und *Moringa oleifera* (*M. oleifera*). *C. spinosa*, *L. douglasii* sowie *M. oleifera* ordnen sich taxonomisch betrachtet in die Ordnung Brassicales ein. Sie gehören jedoch zu den Familien der Cleomaceae, Limnanthaceae und Moringaceae, drei Schwesterfamilien der Brassicaceae (Beilstein et al. 2010). Auch die Spezies *Erysimum cheiri* (*E. cheiri*; Brassicaceae) und *Arabis glabra* (*A. glabra*; Brassicaceae) wurden in den folgenden molekularbiologischen Ansatz einbezogen. Auch für diese Arten war es nicht möglich verlässliche Profile der Glucosinolat-Hydrolyseprodukte zu erhalten (Tabelle 2), sie erschienen jedoch ebenfalls auf Grund ihrer systematischen Stellung interessant und erweiterten die bisher untersuchten Modellorganismen um die Tribus Arabideae und Erysimeae.

Die Klonierung erfolgte mittels DOP-PCR an cDNAs der fünf Spezies mit nachfolgender 5'RACE-PCR und 3'RACE-PCR. Ohne Ergebnis blieben DOP-PCR-Ansätze auf cDNA der Spezies *A. glabra*, *E. cheiri* und *L. douglasii*. Im Verlauf der vorliegenden Arbeit konnte jedoch eine cDNA für ein bisher nicht bekanntes spezifizierendes Protein aus *C. spinosa* isoliert werden, das eine Sequenzidentität von 98 % auf Nukleotidebene (ClustalW2) zu dem in Tabelle 6 aufgeführten ESP aus *I. tinctoria* aufwies. Die Klonierung aus *M. oleifera* erbrachte die cDNA eines spezifizierenden Proteins mit 99 % Sequenzidentität auf Nukleotidebene (ClustalW2) zu dem putativen spezifizierenden Protein aus *C. spinosa*. Auf Grund der extrem hohen Sequenzähnlichkeiten wurden beide dargestellten Klonierungen dreimal unabhängig unter Mitführung von Wasserkontrollen aller Bestandteile des

PCR-Ansatzes und zweimalig mit unabhängig voneinander isolierter RNA wiederholt. Dies erbrachte stets identische Ergebnisse. Zur funktionellen Charakterisierung wurden die ORFs beider cDNAs in *E. coli* zur Expression gebracht. Sowohl für die aus *C. spinosa* als auch die aus *M. oleifera* isolierte cDNA konnte ESP-Aktivität nachgewiesen werden. Auf Grund der enormen Sequenzähnlichkeiten und der daraus resultierenden Vermutung einer eventuellen Verunreinigung wurden die aufgeführten Sequenzen nicht in weitere phylogenetische Untersuchungen einbezogen. Sie werden jedoch an Hand der Tatsache einer dreifach unabhängigen Sequenzbestätigung in der Diskussion berücksichtigt.

2.10.1.1 Suche nach homologen Sequenzen mittels Southern-Kapillar-Blot

Mittels Southern-Kapillar-Blot sollten weitere Hinweise auf das Vorhandensein spezifizierender Proteine in den in Tabelle 10 gelisteten Spezies gesammelt werden.

Tabelle 10: Mittels Southern-Kapillar-Blot analysierte Spezies. Tabellarisch aufgeführt sind nur die Spezies, die nicht in die DOP-PCR der in Tabelle 4 gelisteten Spezies eingegangen sind. Familie und Tribus sind angegeben.

Spezies	Tribus	Familie
<i>Arabis glabra</i>	Arabideae	Brassicaceae
<i>Barbarea intermediata</i>	Cardamineae	Brassicaceae
<i>Barbarea vulgaris</i>	Cardamineae	Brassicaceae
<i>Camelina microcarpa</i>	Camelineae	Brassicaceae
<i>Capsella rubella</i>	Camelineae	Brassicaceae
<i>Cleome spinosa</i>	Cleomeae	Cleomaceae
<i>Eruca sativa</i>	Brassiceae	Brassicaceae
<i>Erysimum cheiri</i>	Erysimeae	Brassicaceae
<i>Limnanthes douglasii</i>	Limnantheae	Limnanthaceae
<i>Moringa oleifera</i>	Moringeae	Moringaceae
<i>Reseda lutea</i>	Resedaeae	Resedaceae
<i>Reseda luteola</i>	Resedaeae	Resedaceae

Für die Herstellung der notwendigen Sonden wurde unter Verwendung des EBI Webservers für multiple Sequenzalignments ClustalW2 (Chenna et al. 2003) ein Multialignment auf Ebene der Aminosäuresequenzen bisher isolierter spezifizierender Proteine erstellt. Durch den Vergleich der ORFs konnten Regionen identifiziert werden, welche sich zur Generation von spezifischen Primern eigneten (Tabelle S 10). Die gewählten Primer flankieren Bereiche der Sequenzen, welche es ermöglichen sollten, durch die entstehenden Sonden ESPs und TFPs getrennt von NSPs nachzuweisen (Abbildung 18). Das Design der Primer erfolgte auf Basis der Nukleotidsequenzen des AtNSP1 (Burow et al. 2009; Kissen and Bones 2009)

sowie des AtESP (Lambrix et al. 2001). Zur Verdeutlichung ist der entsprechende Abschnitt des Alignments in Abbildung 18 dargestellt, das vollständige Alignment sowie die genauen Primersequenzen sind dem Anhang zu entnehmen (Tabelle S 10). Die Sonden wurden unter Verwendung der entsprechenden cDNAs durch Einbau von Digoxigenin-markiertem dUTP (DIG-UTP) generiert. Der erfolgreiche Einbau des DIG-UTP wurde durch Gelelektrophorese des PCR-Ansatzes überprüft. Hierbei weist das markierte DNA-Fragment durch den mehrfachen Einbau des DIG-UTP eine gegenüber der laut Primerpositionen berechneten Größe des unmarkierten DNA-Fragments von 315 bp (AtNSP1) und 300 bp (AtESP) (Abbildung 18) ein erhöhtes Molekulargewicht auf. Erfolgreich generierte Sonden wurden gelelektrophoretisch aufgetrennt, extrahiert und dem Hybridisierungspuffer zugesetzt.

Das spezifische Bindungsverhalten der erhaltenen Sonden und die optimalen Hybridisierungsbedingungen wurden mittels eines Southern-Dot-Blots überprüft (Abbildung 19). Als DNA-Proben wurden Plasmide mit AtNSP1- und AtESP-cDNA punktförmig auf eine Nitrocellulosemembran aufgetragen und nach entsprechender Vorbehandlung mit jeweils einer Sonde bei 43°C und 65°C hybridisiert. Dabei erwiesen sich die Sonden als spezifisch für eine differenzierte Detektion von AtESP und AtNSP1. Das beste Ergebnis für die Detektion von AtESP wurde mit einer Hybridisierungstemperatur von 43°C über Nacht erzielt (Abbildung 19). AtNSP1 konnte bei 43°C sowie 65°C über Nacht detektiert werden (Abbildung 19). Somit hybridisierten die ESP-Sonde bei 43°C ausschließlich mit AtESP, nicht mit AtNSP1 und die NSP-Sonde nur mit AtNSP1, nicht mit AtESP. Auf Grund der hohen Sequenzähnlichkeiten von ESPs und TFPs untereinander sowie NSPs untereinander konnte von einer Übertragbarkeit dieser Ergebnisse auf weitere spezifizierende Proteine ausgegangen werden.

Abbildung 18: Positionen der Primer für die Sonden-Generation. Gezeigt ist der Ausschnitt eines Alignments der Aminosäuresequenzen aller spezifizierenden Proteine. Erstellt wurde das Multialignment mittels des EBI Webservers für multiple Sequenzalignments ClustalW2. Weiße Buchstaben auf schwarzem Grund zeigen Aminosäuren einer Spalte die in mindestens 80 % der Sequenzen identisch sind. Weiße Buchstaben auf grauem Grund zeigen Aminosäuren einer Spalte die in mindestens 80 % der Sequenzen einer Gruppe mit biochemisch ähnlichen Eigenschaften zugeordnet werden können. Positionen und Namensgebung der Primer sind gekennzeichnet: Primer für die NSP-Sonde (blaue Boxen) basieren auf der Nukleotidsequenz des AtNSP1. Primer für die ESP-Sonde (violette Boxen) basieren auf der Nukleotidsequenz des AtESP. Die während der PCR amplifizierten Sondensequenzen sind farblich hinterlegt. NSP-Sonde = blau, ESP-Sonde = violett.

<i>C. hirsuta</i>	ESP	64	IAPANGEAPDVCLGVRMVAVGKLYIFGGRN-EDNHESFYSYDVTVKKEWKLLTKLDDV	
<i>C. impatiens</i>	ESP	64	VVQANGVAPDVCLGVRMVAVGTIIYVFGGRD-ELRHFCDFYSYDVTVKNEWKLLTKLDEA	
<i>A. lyrata</i>	ESP	64	IAPANGDPPTISCLGVRMVAVGKLYIFGGRD-AKFAFNDFYSYDVTQKEWTFLLTKLDEE	
<i>A. thaliana</i>	ESP	64	IAQPKGDAPIVSCLGVRMVAVGKLYIFGGRD-ENRNFEMFYSYDVTSEWTFLLTKLDEV	
<i>B. oleracea</i>	ESP	64	ISPKGVDAPDKALGTRMVAVGKLYIFGGRD-ENKKEFDFYSYDVTNWKLLTKLTDQE	
<i>S. purpurea</i>	ESP	64	IAEVKGDAPIVKA LGTRMVAVGKLYIFGGRD-ENKQDFDFYSYDVTNWKLLTKLTDQE	
<i>I. tinctoria</i>	ESP	64	IAPANGDVNVKALGTRMVAVGKLYIFGGRD-EKQDFDFYSYDVTVKKEWKFLTKLTDQE	
<i>B. rapa</i>	ESP	64	IAPANGQAPNVQALGTRMVAVGTMLYIFGGRD-EKQDFDFYSYDVTVKQEWKFLTKLDEE	
<i>A. petiolata</i>	TFP	64	IAPAKGDVPTKVLGTRMVAVGKLYIFGGRN-KQLEFDFYSYDVTVKKEWKFLTKLDEE	
<i>T. arvense</i>	TaTFP	64	IAPANGDVPKTRVLGTRMVAVGKLYIFGGRN-KQLEFDFYSYDVTVKKEWKFLTKLDEK	
<i>D. aurea</i>	ESP	64	IAQTKGDIPTKLC LGTRMVAVGKLYIFGGRN-ENKDFDFYSYDVTVKNEWKFLTKLDEE	
<i>D. lanceolata</i>	ESP	64	IAQTKGHIPKTKLC LGTRMVAVGKLYIFGGRN-ENKDFDFYSYDVTVKNEWKFLTKLDEE	
<i>L. sativum</i>	TaTFP	64	IAEPKGDVPTKVLGTRMVAVGKLYIFGGRD-ENKNSFYSYDVTNWKLLTKLDEE	
<i>A. thaliana</i>	AtNSP1	207	ISPATGDPVPHLSCLGVRMVSVGSTLYVFGGRD-ASRQYNGFYSFDTTNEWKLLTPVEE-	
<i>A. thaliana</i>	AtNSP4	356	ISPATGDIPTLSCLGVCMSIGSTLYVFGGRD-ASRQYNGFYSFDTTNEWKLLTPVEE-	
<i>A. thaliana</i>	AtNSP3	207	IAPATGDPVPHLSCLGVRMVSVGSTLYVFGGRD-FSRQYNGFYSFDTTNEWKLLTPVEE-	
<i>C. hirsuta</i>	NSP	208	ISPATGDPVPHLSCLGVRMVSVGSNLYVFGGRD-ASRQYNGFYSFDTTNEWKLLTPVEE-	
<i>A. thaliana</i>	AtNSP2	208	ISPATGDPVPHLSCLGVRMVSVGSNLYVFGGRD-ASRQYNGFYSFDTTNEWKLLTPVEE-	
<i>I. tinctoria</i>	NSP	206	IAPATGDPVPHLSCLGVRMVSVGTLNLYVFGGRD-ASRQYNGFYSFDTTNEWKLLTPVEE-	
<i>S. purpurea</i>	NSP	204	ISPATGDPVPHLSCLGVRMVSVGSNLYVFGGRD-ASRQYNGFYSFDTTNEWKLLTPVEE-	
<i>E. halophilum</i>	NSP	63	IAPATGDAP-FPCFVSMVPTIGSTLYVYGGRD-DSRYNGLHSYDTEETNEWKLLAPVEE-	
<i>A. thaliana</i>	AtNSP5	62	IAPATGEAP-FPCFVSMVPTIGSTLYVYGGRD-DKRYNGLHSYDTEETNEWKLLAPVEE-	
<i>A. thaliana</i>	At3g07720	64	IQEASGDAP-PPRVGVMAAVGPIIYFEGGRD-STHQLNLYVCNLTNWKLLSGET-	
				NSP-Sonde fw
<i>C. hirsuta</i>	ESP	123	EAPNNEPAARTYHMASDENHVVYVFGGVSKGGRQTPFRFTAEAYNIATGKWAQLPDP	
<i>C. impatiens</i>	ESP	123	GGPE----ARTYHSMASDQNHVVYVFGGVSKGGKQSTPFRFTTEAYNIAAGKWDTLDPDP	
<i>A. lyrata</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGLQTPKRFRTIEAYNIADGGWVQLPDG	
<i>A. thaliana</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGTMNTPFRRTIEAYNIADGKWAQLPDP	
<i>B. oleracea</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGTNKTPFRFTIEAYNIADGKWSQLPDP	
<i>S. purpurea</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGLMVTTPFRFTIEAYNIADGKWSQLPDP	
<i>I. tinctoria</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGTNKTPFRFTIEAYNIADGKWTQLPDP	
<i>B. rapa</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGTNKTPFRFTIEAYNIADGKWAQLPDP	
<i>A. petiolata</i>	TFP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGLNATPFRFTIEAYNIAEGKWAQLPDP	
<i>T. arvense</i>	TaTFP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGLNATPFRFTIEAYNIAEGKWAQLPDP	
<i>D. aurea</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGKMTTPVRFTIEAYNIADGKWSQLPDP	
<i>D. lanceolata</i>	ESP	123	GGPE----ARTYHSMASDENHVVYVFGGVSKGGKMTTPVRFTIEAYNIADGKWSQLPDP	
<i>L. sativum</i>	TaTFP	123	RVPE----ARSFPAIADQNHVVYVFGGVSKGGVQSTPFRFTIEAYNIAEGKWSQLPDP	
<i>A. thaliana</i>	AtNSP1	265	-GPT----PRSFHSMADENNVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>A. thaliana</i>	AtNSP4	414	-GPT----PRSFHSMADENNVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>A. thaliana</i>	AtNSP3	265	-GPT----PRSFHSMADENNVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>C. hirsuta</i>	NSP	266	-GPT----PRSFHSMADENNVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>A. thaliana</i>	AtNSP2	266	-GPT----PRSFHSMADENNVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>I. tinctoria</i>	NSP	264	-GPI----PRSFHSMADENNVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>S. purpurea</i>	NSP	262	-GPI----PRSFHSMADENNVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>E. halophilum</i>	NSP	120	-GLP----GRSYHSMAGDPRKVVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>A. thaliana</i>	AtNSP5	119	-GLP----GRSYHSMAGDPRKVVYVFGGVSKGGLNATPFRFTIEAYNIADGKWSQLPDP	
<i>A. thaliana</i>	At3g07720	122	-GEP----NRSYHSLTADSQNVYVFGGCG-----VDGRLNDLWAYNVVDQKWKIFPSG	
				ESP-Sonde fw
<i>C. hirsuta</i>	ESP	183	Q---SFEPRGGAGFVEVQGKIWVYGFATSPFPG----GKSNYESDAVHYDFEATKWT	
<i>C. impatiens</i>	ESP	179	Q---SFEPRGGAGFVEVQGKIWVYGFATSPFPG----GKSNYESDAVHYDFEATKWT	
<i>A. lyrata</i>	ESP	179	E---NFEKRGAGFAVQGKIWVYGFATSIIPG----GKNDYESNAQYDFEATKWT	
<i>A. thaliana</i>	ESP	179	D---NFEKRGAGFAVQGKIWVYGFATSIIPG----GKNDYESNAQYDFEATKWT	
<i>B. oleracea</i>	ESP	179	EQFPRFERRGAGFVQVQGIWVYGFATSPDP----NGKNDYESDQVQYDFEATKWT	
<i>S. purpurea</i>	ESP	179	EQFPRFERRGAGFVQVQGIWVYGFATSPDP----NGKNDYESDQVQYDFEATKWT	
<i>I. tinctoria</i>	ESP	179	EQFHNFEKRGAGFVQVQGIWVYGFATSPDP----NGKNDYESDQVQYDFEATKWT	
<i>B. rapa</i>	ESP	179	EQFPRFERRGAGFVQVQGIWVYGFATSPDP----NGKNDYESDQVQYDFEATKWT	
<i>A. petiolata</i>	TFP	179	E---DFEKRGMAGFLVQVQGIWVYGFATANDPKIPTLYGSQDYESNRVHCFEATKWT	
<i>T. arvense</i>	TaTFP	179	E---DFEKRGMAGFLVQVQGIWVYGFATANDPKIPTLYGSQDYESNRVHCFEATKWT	
<i>D. aurea</i>	ESP	179	E---EFEKRGAGFHVQGRVWVYGFATSIIEVG----GKNDYESNHYQYDFEATKWT	
<i>D. lanceolata</i>	ESP	179	E---EFEKRGAGFHVQGRVWVYGFATSTQVE----GKNDYESNHYQYDFEATKWT	
<i>L. sativum</i>	TaTFP	179	P---DFEPRGGAGLAVDKKLWVYGFANSTSG----GINDYNSNKVQYDFEATKWT	
<i>A. thaliana</i>	AtNSP1	314	D---SLTARGGAGLEVQVQGIWVYGFANSTSG----GCEVDDVHYDFEATKWT	
<i>A. thaliana</i>	AtNSP4	463	D---SLTARGGAGLEVQVQGIWVYGFANSTSG----GCEVDDVHYDFEATKWT	
<i>A. thaliana</i>	AtNSP3	314	D---SFSIRGGAGLEVQVQGIWVYGFANSTSG----GCEVDDVHYDFEATKWT	
<i>C. hirsuta</i>	NSP	315	D---YFSIRGGAGLEVQVQGIWVYGFANSTSG----GCEVDDVHYDFEATKWT	
<i>A. thaliana</i>	AtNSP2	315	G---SCSVRGAGLEVQVQGIWVYGFANSTSG----GCEVDDVHYDFEATKWT	
<i>I. tinctoria</i>	NSP	313	E---ALSIRGGAGLEVQVQGIWVYGFANSTSG----GQETDDVHYDFEATKWT	
<i>S. purpurea</i>	NSP	311	E---SVSIRGGAGLEVQVQGIWVYGFANSTSG----GCEVDDVHYDFEATKWT	
<i>E. halophilum</i>	NSP	169	E---ACKGRGAPGLVQVQGIWVYGFANSTSG----GNELGDFHCFEATKWT	
<i>A. thaliana</i>	AtNSP5	168	E---ACKGRGAPGLVQVQGIWVYGFANSTSG----GNELGDFHCFEATKWT	
<i>A. thaliana</i>	At3g07720	171	E---ACRGRGAPGLVQVQGIWVYGFANSTSG----GEEADDVHYDFEATKWT	
				ESP-Sonde rev
				NSP-Sonde rev

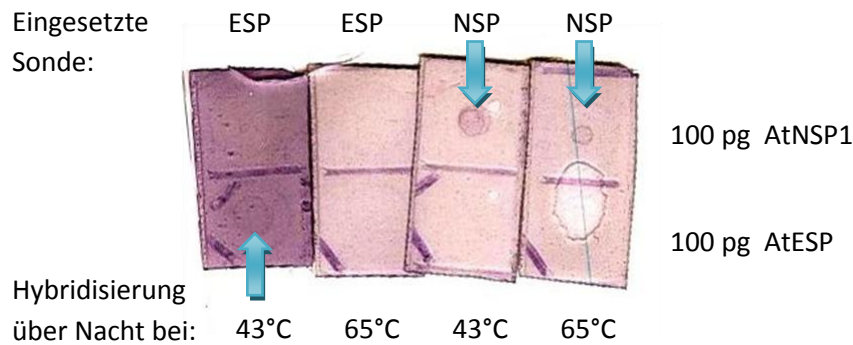


Abbildung 19: Southern-Dot-Blot zur Bestimmung spezifischer Hybridisierungsbedingungen. Gezeigt wird ein Dot-Blot von Plasmid-DNA (je 100 pg cDNA von AtNSP1 und AtESP). Die Hybridisierung mit der NSP-Sonde und der ESP-Sonde erfolgte separat bei 43°C oder 65°C über Nacht. Zur Detektion wurde das Enzym alkalische Phosphatase unter Verwendung von NBT und BCIP als Farbkomplexbildner verwendet. Spezifisch detektierte Spots sind durch Pfeile gekennzeichnet.

Im zeitlichen Rahmen der vorliegenden Arbeit gelang es trotz Anwendung verschiedener Methoden nicht, ausreichend genomische DNA zur direkten Verwendung in einem für den Southern-Kapillar-Blot notwendigen Restriktionsverdau zu isolieren. Um dessen ungeachtet eine Beantwortung der oben aufgeführten Fragestellung zu erzielen, wurde alternativ eine PCR mit der isolierten genomischen DNA als Template durchgeführt. Zum Einsatz kamen alle in Tabelle 6 und Tabelle 10 aufgeführten Spezies sowie eine Auswahl vorhandener einfacher und degenerierter Primer (siehe Abbildung 20 sowie Tabelle S 10). Erhaltene PCR-Fragmente wurden mittels Agarose-Gelelektrophorese analysiert. Die Größe theoretisch zu erwartender DNA-Fragmente wurden an Hand aller bekannten Sequenzen spezifizierender Proteine bestimmt und lag in einer Größenordnung von 672 bis 798 Nukleotiden unter Verwendung der degenerierten Primer und bis zu 1380 Nukleotiden bei Einsatz der Primerkombination P149 und P150. Die Verwendung aller weiteren Primerkombinationen führte zu einer theoretisch berechneten Größe der DNA-Fragmente von 750 bis 1000 Nukleotiden. Das erhaltene Agarose-Gel (Abbildung 20A) wurde direkt für einen Southern-Kapillar-Blot verwendet. Zur Hybridisierung der Membran wurde in einem ersten Durchlauf die ESP-Sonde zu Detektion von ESPs und TFPs verwendet (Abbildung 18, Abbildung 20B). Nach Stripping der Membran wurde diese sofort erneut prähybridisiert und in einem zweiten Durchlauf einer Hybridisierung mit der NSP-Sonde zum Nachweis von NSPs unterzogen (Abbildung 18, Abbildung 20C).

Das in Abbildung 20A gezeigte Bild des entsprechenden Agarose-Gels zeigt deutlich die Vielzahl erhaltener DNA-Fragmente in einer Größenordnung von < 200 bp - ~ 1000 bp für die einzelnen Spezies und Primerkombinationen. Ausnahmen bilden die Spezies *A. glabra* und *Barbarea vulgaris* (*B. vulgaris*) ohne jegliche Banden.

Abbildung 20B zeigt den Blot nach Hybridisierung mit der ESP-Sonde. In den Spalten 1 – 14, sind deutlich Banden auf Höhe der zuvor im Agarose-Gel nachgewiesenen und unter Verwendung der degenerierten P1 und P2 – zur Amplifikation von ESPs und TFPs - erhaltenen DNA-Fragmente zu erkennen. Zudem konnten auf gleicher Höhe Banden für die Spezies *A. glabra* und *B. vulgaris* erhalten werden (Spalten 1 und 2). Als einzige Spezies ohne Bande stellte sich *M. oleifera* (Spalte 11) dar. Die unter Verwendung der Primer P5 und

P6 – zur Amplifikation von NSPs- gewonnenen DNA-Fragmente in den Spalten 15 – 28 zeigten mit Ausnahme von Schattenbanden bei *Eruca sativa* (*E. sativa*) (Spalte 22) und *I. tinctoria* (Spalte 24) keine Banden. Dieses Ergebnis bestätigt die Spezifität der generierten ESP-Sonde. In den Spalten 29 – 58 wurden PCR-Ansätze aufgetragen, welche unter Verwendung von Kombinationen aus degenerierten und einfachen Primern sowie unter Verwendung bereits bei der Identifizierung von spezifizierenden Proteinen verwendeter Primerkombinationen durchgeführt wurden. Zum Einsatz kamen neben *I. amara* weitere Spezies, welche bereits im Vorfeld die Identifizierung putativer (2.10.1) und bereits charakterisierter (2.4) spezifizierender Proteine ermöglichten. Wurden dieselben Hybridisierungsbedingungen auf diesen PCR-Reaktionen angewendet, konnten ausschließlich Ergebnisse für PCR-Amplifikate unter Verwendung von Primerkombinationen zur Klonierung von ESPs in den Spezies *C. spinosa* (Spalten 29, 31, 33), *I. amara* (Spalten 39 – 43) und *I. tinctoria* (Spalten 49 – 54) erhalten werden. Dieses Ergebnis bestätigt erneut die Spezifität der eingesetzten ESP-Sonde. Somit konnten für PCR-Reaktionen auf genomischer DNA unter Verwendung von Primerkombinationen für eine Amplifikation von *ESP*- und *TFP*-Sequenzen entsprechend ähnliche Sequenzen im Genom der untersuchten Spezies außer für *M. oleifera* spezifisch nachgewiesen werden.

Abbildung 20C zeigt den Blot nach Stripping und erneuter Hybridisierung der Membran mit der NSP-Sonde. In den Spalten 4 – 14 ergab sich eine deutliche Kreuzreaktion mit der ersten Hybridisierung mittels der erstellten ESP-Sonde. Die erkennbaren Banden basieren auf DNA-Fragmenten, welche mit den degenerierten Primern P1 und P2 amplifiziert wurden. In den Spalten 15 – 28 der unter Verwendung der degenerierten Primer P5 und P6 durchgeführten PCR-Reaktionen konnten jedoch Banden für *B. oleracea*, *Camelina microcarpa* (*C. microcarpa*), *E. cheiri*, *E. sativa*, *I. amara*, *I. tinctoria* und *Reseda lutea* (*R. lutea*) detektiert werden. Da die Primer P5 und P6 für die Amplifikation von putativen NSPs generiert wurden, kann aus diesem Ergebnis auf eine überwiegend spezifische Reaktion der Sonde geschlossen werden. Diese Aussage lässt sich durch das Fehlen von Banden in den Spalten 29 – 43 sowie Spalten 49 – 53 unterstützen, da hier erwartete DNA-Fragmente auf Grund der für die PCR-Reaktion eingesetzten Spezies und Primerkombinationen eine hohe Ähnlichkeit zu ESPs aufweisen sollten. In den Spalten erwarteter NSP-Amplifikate der Spezies *I. tinctoria* (Spalten 54 - 58) konnten zudem Banden detektiert werden. Somit konnten *NSP*-Sequenzen im Genom der aufgeführten Spezies spezifisch für DNA-Fragmente aus PCR-Reaktionen unter Verwendung von Primerkombinationen zur Amplifikation von NSPs nachgewiesen werden. Dieser Nachweis gelang nicht für *NSP*-ähnliche Sequenzen in den Spezies *A. glabra*, *B. vulgaris*, *Capsella rubella* (*C. rubella*), *C. spinosa*, *M. oleifera*, *R. lutea*, *Reseda luteola* (*R. luteola*) und *T. minus*.

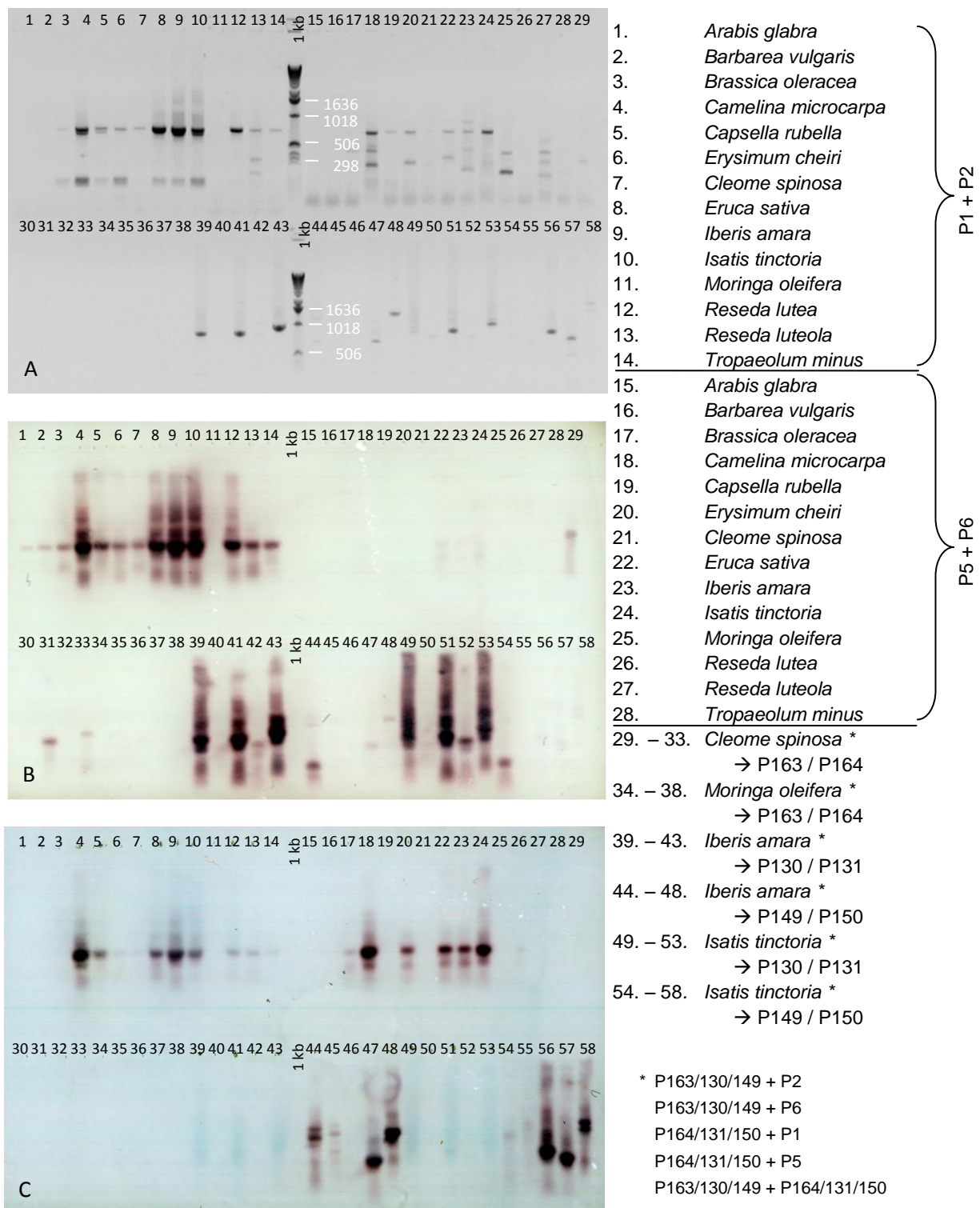


Abbildung 20: Southern-Kapillar-Blot-Analyse von PCR-Reaktionen auf genomischer DNA mit ESP- und NSP-spezifischen Sonden. (A) Abgebildet ist das Lichtbild eines Agarose-Gels. Banden wurden durch den Einsatz verschiedener Primerkombinationen (Legende rechts) auf genomischer DNA der aufgeführten Spezies als Template erhalten und durch Verwendung von Ethidiumbromid im UV-Licht detektiert. Aufgetragen wurde zudem ein 1 kb Leiter (mittig) zum Größenvergleich erhaltener Banden. Das erhaltene Agarose-Gel wurde direkt für einen Southern-Kapillar-Blot verwendet. (B) Southern-Kapillar-Blot unter Verwendung der ESP-Sonde. (C) Southern-Kapillar-Blot unter Verwendung der NSP-Sonde. Spezies sowie die eingesetzte Primerkombination sind der Nummerierung zu entnehmen. Die Primerkombination P1 mit P2 wurden für die Amplifikation ESP / TFP-ähnlicher Fragmente und die Primerkombination P5 mit P6 für die Amplifikation NSP-ähnlicher Fragmente verwendet.

2.11 Charakterisierung des Thiocyanat-formenden Proteins aus *T. arvense*

Das TaTFP (Kuchernig et al. 2011) ist das erste auf molekularer Ebene identifizierte TFP, das die Bildung von organischem Thiocyanat bei der Myrosinase-katalysierten Hydrolyse von Allylglucosinolat unterstützt. Aus diesem Grund sollte es genauer untersucht werden. Dabei sollte im Rahmen der vorliegenden Arbeit das Expressionsmuster von TaTFP in der Pflanze dem Glucosinolatprofil und dem Profil der Glucosinolat-Hydrolyseprodukte der jeweiligen Pflanzenorgane gegenübergestellt werden. Im Rahmen der vorliegenden Arbeit wurden die Glucosinolatprofile sowohl von ober- als auch unterirdischen Pflanzengewebe sowie Keimlingen dreimal unabhängig voneinander bestimmt.

Hauptglucosinolat in Keimling, Stängel, Blatt, Blüte, Schote und Samen war mit bis zu 99,7 % Anteil am Gesamtglucosinolatgehalt das Allylglucosinolat (Abbildung 21, Tabelle S 9). Benzylglucosinolat konnte mit einem deutlich geringeren Gehalt von durchschnittlich max. 2 % (Keimling) dokumentiert werden (Abbildung 21, Tabelle S 9). Weitere Glucosinolate traten nur in geringeren Mengen oder in Spuren auf (Abbildung 21, Tabelle S 9). In den reproduktiven Geweben (Blüte, Schote, Samen) ist deutlich der vergleichsweise hohe Gesamtglucosinolatgehalt mit bis zu durchschnittlich 26 $\mu\text{mol} / 100 \text{ mg}$ Frischgewicht in Samen, bei eingeschränkter Vielfalt der auftretenden Glucosinolate zu erkennen (Abbildung 21, Tabelle S 9). Bemerkenswert ist das entsprechend breite Spektrum an Glucosinolaten in der Wurzel. Ebenfalls in der Wurzel wurde der geringste Glucosinolatgehalt nachgewiesen (1,12 $\mu\text{mol} / 100 \text{ mg}$ Frischgewicht) (Tabelle S 9).

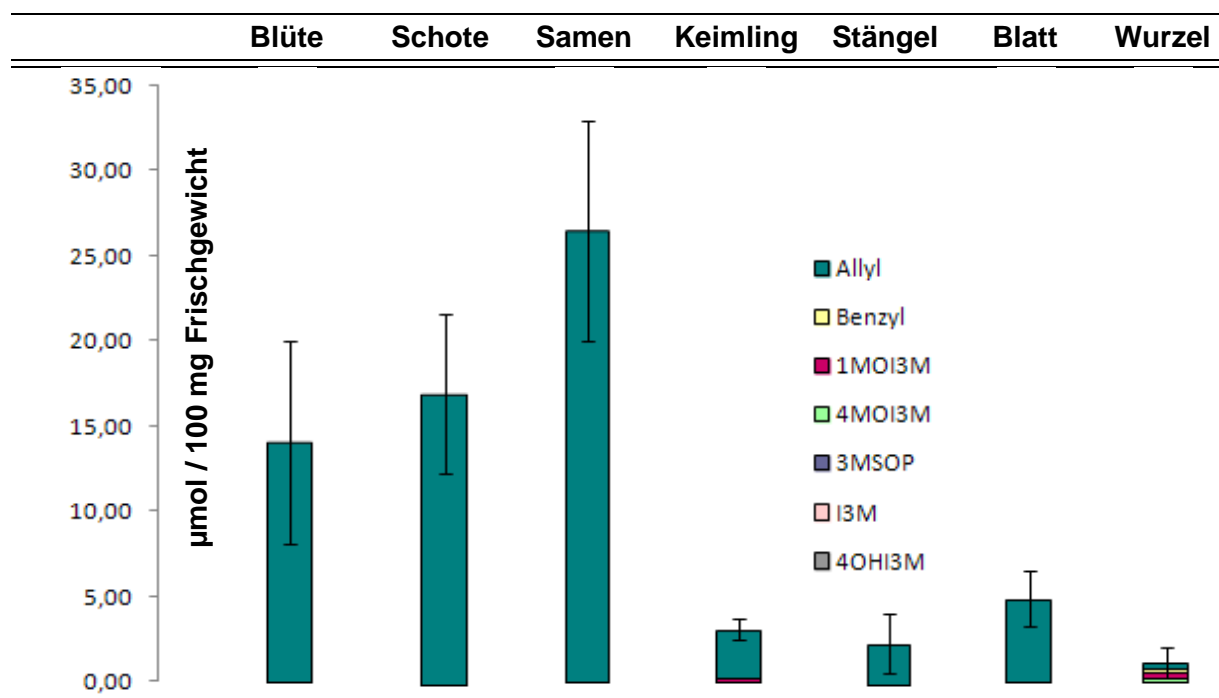


Abbildung 21: Glucosinolatgehalt in *T. arvense*. Der Gehalt an Glucosinolaten mit folgenden Seitenketten wurde durch HPLC der Desulfoglucosinolate bestimmt: Allyl, Benzyl, Indol-3-yl-methyl (I3M), 1-Methoxy-indol-3-ylmethyl (1MOI3M), 4-Methoxy-indol-3-ylmethyl (4MOI3M), 4-Hydroxy-indol-3-ylmethyl (4OHI3M) und 3-Methylsulfinylpropyl (3MSOP). Angegeben sind die Mittelwerte aus drei unabhängigen Versuchen. In jedem Versuch wurden ein bis sechs Individuen analysiert. Die Daten sind Tabelle S 9 im Anhang zu entnehmen.

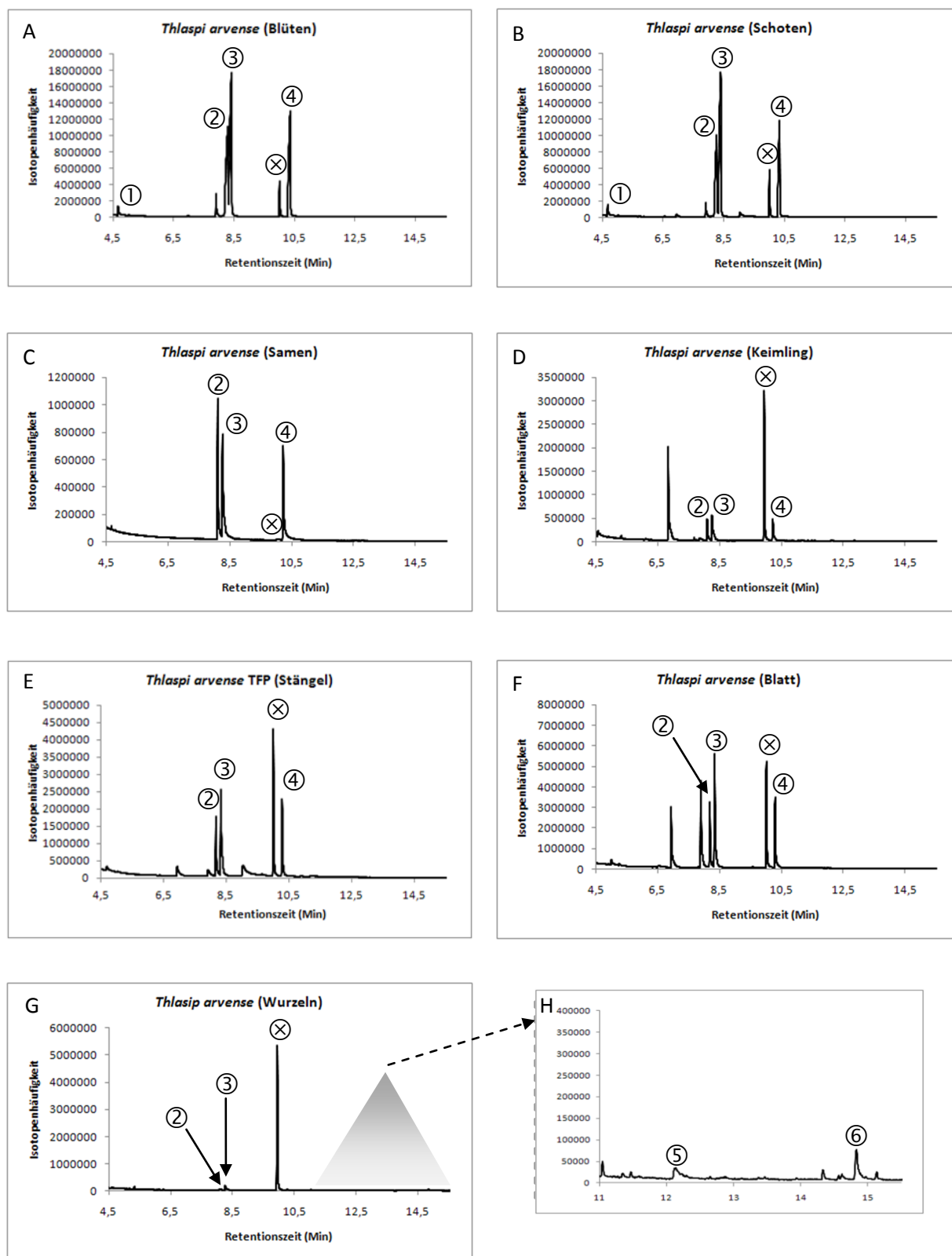


Abbildung 22: Organspezifische Bestimmung der Glucosinolat-Hydrolyseprodukte in *T. arvense*. Gezeigt werden repräsentative GC-MS-Chromatogramme von Dichlormethanextrakten von Autolysaten aus Blüten (A), Schoten (B), Samen (C), Keimlingen (D), Stängeln (E), Blättern (F) und Wurzeln (G, H) aus *T. arvense*. Nicht markierte Peaks entsprechen nicht relevanten Hintergrundprodukten, welche der abgelaufenen Myrosinase-Reaktion nicht zuzuordnen sind. Weitere Peaks entsprechen: ① = einfaches Nitril aus Allylglucosinolat, ② = Thiocyanat aus Allylglucosinolat, ③ = Isothiocyanat aus Allylglucosinolat, ④ = Epithionitril aus Allylglucosinolat, ⑤ = einfaches Nitril aus Benzylglucosinolat, ⑥ = Isothiocyanat aus Benzylglucosinolat, ⊗ = interner Standard Phenylcyanid

Um den Umsatz der identifizierten Glucosinolate nach Ablauf der Glucosinolat-Myrosinase-Reaktion zu untersuchen, wurden Dichlormethanextrakte aus Autolysaten verschiedener Organe per GC-MS analysiert (Abbildung 22). Isothiocyanat sowie Thiocyanat und Epithionitril aus Allylglucosinolat konnten in Samen, Blüten, Stängeln, Blättern, unreifen Schoten sowie zwei Wochen alten Keimlingen nachgewiesen werden (Abbildung 22A - F). In den Autolysaten aus Wurzeln konnte bei Einsatz der gleichen Menge Pflanzenmaterial auf Grund des sehr geringen Gehalts an Allylglucosinolat kein Epithionitril aus Allylglucosinolat nachgewiesen werden (Abbildung 22G + H). Durch den in diesem Gewebe erhöhten Anteil an Benzylglucosinolat war es jedoch möglich, Spuren von einfachem Nitril sowie Isothiocyanat aus Benzylglucosinolat zu dokumentieren. Weiterhin nachweisbar war das einfache Nitril aus Allylglucosinolat in Blüten und unreifen Schoten sowie Samen. Auf Grund der sehr geringen Mengen weiterer möglicher Substrate bei sehr großen Mengen Allylglucosinolat in generativen Organen (Abbildung 21, Tabelle S 9) war es nicht möglich, größere Mengen zu analysieren, da dies zu einer Überladung der GC-MS geführt hätte.

Zur Expressionanalyse von TaTFP mittels semiquantitativer PCR (5.2.1.2.2, 5.2.1.3.2) wurden genspezifische Primer basierend auf der cDNA des TaTFP entworfen (Abbildung 23, Tabelle S 10), die ein Amplifikat der berechneten Größe von 415 bp ergaben. Die Identität der erhaltenen DNA-Fragmente wurde durch Sequenzierung abgesichert. Als Kontrolle wurden Primer für die Amplifikation eines 412 bp großen DNA-Fragmentes von *Actin8* (At1g49240, AAC49523, (An et al. 1996)) verwendet (Tabelle S 1). Die Möglichkeit zur Generation der Primer gemäß der *Actin8*-Sequenz aus *A. thaliana* für den Einsatz auf cDNA aus *T. arvense* ergibt sich aus der Sequenzähnlichkeit der beiden Spezies von 90 % - 92 % (Sharma et al. 2007). Die Funktionalität der Primer wurde durch Sequenzierung der erhaltenen DNA-Fragmente überprüft.

```

5' atggctcggactttgcaaggcgagtggtgaaggtggagcagaaaaggaggacaagtaccagcaccgagaag
ctcacacggcatagccgtgatcgagacaagctctattgtttcggcggcgaggatccgccatatgagtccatc
gacaacgacctttatgtctttgacttcaacacccacacttggtcaatcgctccggccaacggagacgttccaa
agaccagagtcttaggcacccgcatggtggccgtgggaacgaagctctatgtgttcggaggccgcaataaaca
gtagagttcggaggacttttactcgtacgatacggtgaaagaggagtggaaagttcctgacgaagctggacgaa
aagggaggacccgaagctcgtactttccattcgatgacttcagatgaaaaccacgtgtacgtgttcggtgggg
tgagcaaaggagggtgaatgcaaccccccttcggttcaggacgatcgaggcctataacattgctgaagggaa
atgggctcagctccctgaccttgagagatttcgagaaaagaggaatggccggattccttgtggtgcaaggc
aagctttgggtgttttacggattcgcgactgcgaatgatcctaagataccacactctacgggtcccaggact
acgagtctaatacgtgtgcactgctatgatcccgctactcaaaaatggaccgaagtggagaccacaggtttcga
gaaaccttctcgtaggagctgttttgccgatcgcgctgtggggaaatatataataatttcggaggtgagatt
gagcgggaccacagaagcacatcaaggtccggggacgttgtcccgtaggggttttgcgttgacactgagacat
tgggtgtgggagaggtatgaaggaggaccaatcaaacggagcaaccgcggttggtggcctccacgacgaccac
catcaacggaaagaaaggtctgctcgtgcgtgaggggaaacttatgaccaacgagcgcaccgatgagatgtac
ttcttcgcagtcgaattcctccacgtaa 3'

```

Abbildung 23: Position der Primer für die semiquantitative PCR in *T. arvense*. Darstellung der Nukleotidsequenz (vollständiger ORF) des TaTFP unter Angabe des für die semiquantitative PCR verwendeten Primerpaares. Positionen, Namensgebungen und Sequenzen sind gekennzeichnet durch blaue Boxen. Die exakten Primersequenzen sind zudem Tabelle S 10 zu entnehmen.

Die aus der semiquantitativen PCR hervorgegangenen Ergebnisse zeigten ein einheitliches Expressionslevel des TaTFP in allen betrachteten ober- und unterirdischen Pflanzengeweben sowie generativen Organen (Wurzeln, Stängeln, Blättern, Blüten, unreifen Schoten sowie zwei Wochen alten Keimlingen) (Abbildung 24). Nicht dargestellt wurden die Ergebnisse aus Samen, da eine eindeutige Identifizierung des PCR-Fragments hier nicht gelungen ist. Neben der *Actin8*-Kontrolle wurde als „Ladekontrolle“ der semiquantitativen PCR jeweils auch 1 µg der für die cDNA-Synthese eingesetzten RNA elektrophoretisch analysiert und der Einsatz einheitlicher RNA-Mengen nachweisen werden (Abbildung 24).

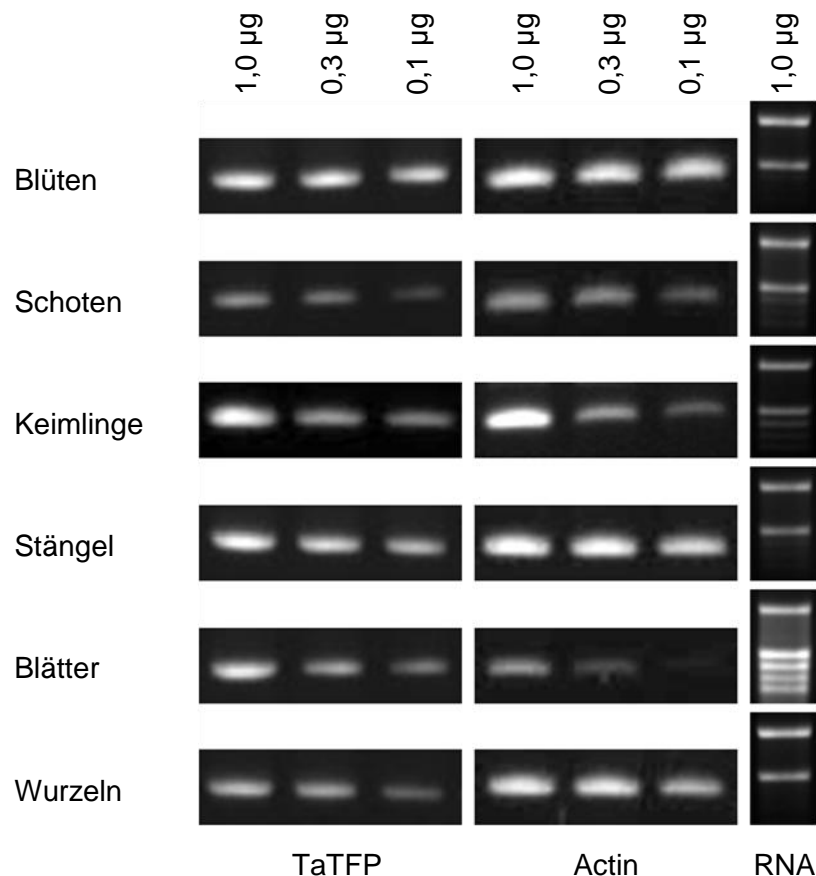


Abbildung 24: Expressionsmuster von TaTFP. Transkriptlevel wurden bestimmt durch den Einsatz genspezifischer Primer (Tabelle S 10) auf cDNA transkribiert aus je 1 µg, 0,3 µg und 0,1 µg RNA ober- und unterirdischer Organe (links) von *T. arvense*. Als Kontrolle mitgeführt wurden semiquantitative PCR-Ansätze auf der jeweils selben cDNA-Präparation unter Einsatz von *Actin8*-spezifischen Primern. Je 27 µl der semiquantitativen PCR-Ansätze sowie wurden mittels Agarose-Gelelektrophorese analysiert und Banden unter Verwendung von Ethidiumbromid im UV-Licht detektiert. Gezeigt ist das Lichtbild eines Agarosegels. Als weitere Kontrolle wurde je 1 µg der RNA jedes elektrophoretisch analysiert (rechts). Das gesamte Experiment wurde insgesamt zwei Mal mit unabhängigen Aussaaten von *T. arvense* wiederholt.

3. DISKUSSION

3.1 Klonierung und Charakterisierung spezifizierender Proteine

Die Zielsetzung der vorliegenden Arbeit lag in der Erarbeitung eines tieferen Einblickes in die evolutive Abstammung der spezifizierenden Proteine des Glucosinolat-Myrosinase-Systems.

Für eine aussagekräftige Phylogenie ist eine möglichst große Anzahl repräsentativer Sequenzen bekannter Funktion der zu betrachtenden Genfamilie zu gewinnen. Hierbei ist es wünschenswert, eine große Bandbreite verschiedener Spezies abzudecken. Für Hinweise auf das Vorhandensein von spezifizierenden Proteinen wurde ein phytochemisches Screening von Spezies der Ordnung Brassicales durchgeführt, welches die Grundlage für die Auswahl geeigneter Spezies zur Isolierung spezifizierender Proteine bildete (2.1). Im phytochemischen Screening wurden Spezies aus sechs Familien (Brassicaceae, Cleomaceae, Limnanthaceae, Moringaceae, Resedaceae, Tropaeolaceae) untersucht. Innerhalb der Familie der Brassicaceae konnten hierbei Spezies aus neun Tribus der Abstammungslinien I und II der Familie Brassicaceae (Al-Shehbaz et al. 2006; Beilstein et al. 2006) (Abstammungslinie I: Camelinaeae, Cardamineae, Erysimeae; Abstammungslinie II: Alysseae, Arabideae, Brassiceae, Iberideae, Isatideae, Thlaspidiae) untersucht werden (Tabelle 2, Tabelle S 1). Innerhalb dieser 30 Spezies konnten 14 Spezies der Familie der Brassicaceae (Tribus Cardamineae der Abstammungslinien I und Alysseae, Arabideae, Brassiceae, Iberideae, Isatideae, Thlaspidiae der Abstammungslinie II) sowie eine Spezies in der Familie Tropaeolaceae (*T. minus*) identifiziert werden, welche von Isothiocyanaten abweichende Glucosinolat-Hydrolyseprodukte bildeten. In *T. minus* wurden jedoch nur sehr geringe Mengen an einfachen Nitrilen gefunden. Da einfache Nitrile von Myrosinasen bei pH-Werten unter fünf und bei hohen Eisenionenkonzentrationen auch ohne den Einfluss spezifizierender Proteine entstehen können (Ettlinger et al. 1961; Tookey and Wolff 1970; Burow et al. 2006a) ist die Bildung einfacher Nitrile alleine kein Beweis für das Vorhandensein spezifizierender Proteine.

Für das phytochemische Screening innerhalb der Brassicaceae waren ausschließlich Spezies der Abstammungslinien I und II verfügbar. Die Untersuchung von Spezies aus der Abstammungslinie III (Al-Shehbaz et al. 2006; Beilstein et al. 2006) wäre ein Gewinn für die Erforschung der phylogenetischen Entwicklung spezifizierender Proteine. Auch Untersuchungen an Spezies des Tribus Aethimoneae könnten aussagekräftige Ergebnisse liefern. Gleichermäßen interessant wäre eine ergänzende Untersuchung der Gattung *Drypetes* aus der Familie der Putranjivaceae. Es handelt sich dabei um die einzige Pflanzengattung außerhalb der Ordnung der Brassicales, in welcher Glucosinolate nachgewiesen werden konnten. Vermutet wird hier ein unabhängiger Ursprung der Glucosinolat-Biosynthese (Rodman et al. 1998).

Die für eine aussagekräftige phylogenetische Analyse der evolutiven Zusammenhänge notwendige repräsentative Anzahl spezifizierender Proteine wurde durch die Identifizierung neuer spezifizierender Proteine (2.3) und ihre funktionelle Charakterisierung (2.4) sowie eine

Datenbankrecherche erreicht (2.5). Für die Klonierung der spezifizierenden Proteine wurde ein DOP-PCR-Ansatz mit degenerierten Primern gewählt (2.2). Degenerierte Primer stellen eine Mischung aus ähnlichen, jedoch divergenten Primern dar. Sie dienen der Identifizierung gleichartiger, jedoch wahrscheinlich nicht identischer Gene aus unterschiedlichen Organismen, deren exakte Sequenz bisher nicht bekannt war. Das Design der Primer basierte hierbei nicht auf einer allgemeinen Codon-Usage, sondern auf den Nukleotidsequenzen der bekannten spezifizierenden Proteine. Daher wurden nach der Identifizierung konservierter Regionen innerhalb der Aminosäuresequenzen der bis zu diesem Zeitpunkt bekannten acht spezifizierenden Proteine die Nukleotidsequenzen der NSPs getrennt von denjenigen der ESPs und TFPs für das abschließende Primerdesign betrachtet (2.2). Der Einsatz degenerierter Primer ermöglichte es, Sequenzvariationen der verschiedenen spezifizierenden Proteine zu erfassen, dabei jedoch in den meisten Fällen eine Differenzierung von *NSP*- und *ESP* / *TFP*-Sequenzen zu erhalten. Unter Einsatz des Primerpaares P1 und P2 konnten überwiegend cDNA-Fragmente putativer *ESP*- und *TFP*-Sequenzen identifiziert werden, unter Einsatz des Primerpaares P5 und P6 fast ausschließlich cDNA-Fragmente putativer *NSP*-Sequenzen (Tabelle 5). Alle mittels DOP-PCR-Ansatz gewonnenen DNA-Fragmente bildeten dabei mittig gelegene Abschnitte der gesuchten cDNAs mit einer Größe von 721 bis 733 Nukleotiden (P1 und P2) und 670 bis 679 Nukleotiden (P5 und P6) (Abbildung 8). Ein Großteil der erhaltenen cDNA-Fragmente konnte durch 3'RACE- und 5'RACE-PCR auf die volle Länge einschließlich 3' und 5'UTR erweitert werden (2.3). Chimäre Artefakte wurden durch unabhängige Amplifikation der ORFs aus Gesamt-cDNA ausgeschlossen. *T. minus* sowie auch die Spezies *A. alpestre* konnten auf Grund zu gering vorhandener Samenmengen nicht für die DOP-PCR eingesetzt werden.

Die funktionelle Zuordnung der ORFs erfolgte durch anschließende Analyse der heterolog exprimierten Proteine (2.4). Im Rahmen der vorliegenden Arbeit konnten ausgehend von acht im Vorfeld bekannten spezifizierenden Proteinen (fünf NSPs, zwei ESPs, ein TFP) (Tabelle 1) und der homologen Sequenz At3g07720 (Burow et al. 2009; Kissen and Bones 2009) elf weitere spezifizierende Proteine aus acht Spezies der Familie Brassicaceae identifiziert und charakterisiert werden (drei NSPs: *C. hirsuta*, *I. tinctoria*, *S. purpurea*, sechs ESPs: *C. hirsuta*, *C. impatiens*, *D. aurea*, *D. lanceolata*, *I. tinctoria*, *S. purpurea*, zwei TFPs: *A. petiolata* und *T. arvense*) (Tabelle 6). Somit konnte die Auswahl zur Verfügung stehender charakterisierter spezifizierender Proteinen von den drei Tribus Camelinae, Brassiceae und Lepidieae auf die zusätzlichen Tribus Thlaspidiae, Cardamineae, Arabideae und Isatideae erweitert werden. Dabei gelang es in drei Fällen spezifizierende Proteine unterschiedlicher Typen aus derselben Art zu isolieren, so dass insgesamt vier paraloge Paare für die phylogenetischen Analysen zur Verfügung standen (*A. thaliana*, *C. hirsuta*, *I. tinctoria*, *S. purpurea*). Unter Einbeziehung von drei einer Datenbankrecherche entnommenen putativen spezifizierenden Proteinen (ein NSP: *E. halophilum*, zwei ESPs: *B. rapa*, *A. lyrata*) (Tabelle 7) konnte eine detaillierte phylogenetische Analyse auf Basis von 22 cDNAs spezifizierender Proteine sowie der ähnlichen Sequenz At3g07720 durchgeführt werden, die acht Tribus der zwei Abstammungslinien I und II der Familie Brassicaceae (Al-Shehbaz et al. 2006; Beilstein et al. 2006) umfasste. Zusätzlich konnten sechs weitere cDNAs fragmentarisch aus den Spezies *A. petiolata*, *C. hirsuta*, *D. cinerea* und *E. hieraciifolium*

isoliert werden (Tabelle 5). Hier gelang es nicht, die DOP-PCR Fragmente mittels RACE-PCR zu vervollständigen. In einigen Fällen stellte sich die RNA-Isolation als problematisch dar. Die Proben wiesen jeweils eine deutliche Trübung auf, welche einen Hinweis auf vorhandene sekundäre Pflanzenstoffe und / oder Phenole lieferte. Diese könnten die folgenden enzymatischen Reaktionen inhibiert haben. Auch DNA-Sekundärstrukturen könnten eine erfolgreiche PCR-Reaktion gestört haben. Die nur partiell isolierten Sequenzen wurden nicht in die phylogenetischen Analysen einbezogen, da eine funktionelle Zuordnung der spezifizierenden Proteine nicht sicher möglich war. Einige Ansätze der DOP-PCR blieben zudem ohne Erfolg (2.3).

Trotz der fehlenden Verfügbarkeit von spezifizierenden Proteinen aus Spezies der Abstammungslinien III und der Tatsache, dass nur eine Sequenz mit Ähnlichkeit zu AtNSP5 zur Verfügung stand und zu TFPs paraloge ESPs oder NSPs fehlen, konnte durch die Verfügbarkeit von insgesamt 23 cDNAs aus einer umfangreichen Artenauswahl und der breiten Fächerung an Spezies aus verschiedenen Tribus innerhalb der Abstammungslinien I und II ein aussagekräftiges Ergebnis der phylogenetischen Analysen erwartet werden. Die Identifizierung weiterer Sequenzen mit Ähnlichkeit zu AtNSP5 wäre für die Erweiterung der berechneten Phylogenie erstrebenswert, da sich alle weiteren identifizierten *NSP*-Sequenzen in ihrer Ähnlichkeit AtNSP1 bis AtNSP4 zuordneten. Ein weiterer wichtiger Punkt wäre die Identifizierung von weiteren TFP-codierenden Sequenzen um ihr Verwandtschaftsverhältnis zu ESPs und NSPs zu festigen. Hierzu notwendig wären zudem zu TFPs paraloge *NSP*- und *ESP*-Sequenzen, um Berechnungen für Vergleiche zwischen orthologen und paralogenen Sequenzen erstellen zu können. Auch die Identifizierung weiterer Sequenzen mit Ähnlichkeit zu At3g07720 könnte die Basis der erstellten Phylogenie untermauern. Eine Isolierung von zu At3g07720 homologen Sequenzen war jedoch nicht Ziel der vorliegenden Arbeit und wurde in der gewählten Klonierungsstrategie nicht berücksichtigt.

In Arten außerhalb der Familie der Brassicaceae wurden bisher keine von Isothiocyanaten abweichenden Glucosinolat-Hydrolyseprodukte identifiziert (eine Ausnahme bildet *T. minus* (2.1)). Die Glucosinolatprofile und die bei Gewebeverletzung freigesetzten Glucosinolat-Hydrolyseprodukte hängen jedoch von Pflanzenorgan, Entwicklungszustand, Wachstumsbedingungen und Befall mit Herbivoren und Pathogenen ab (Brown et al. 2003; Charron et al. 2005; Charron et al. 2005; Himanen et al. 2008; Wentzell and Kliebenstein 2008; Textor and Gershenzon 2009). Daher würden die untersuchten Spezies möglicherweise unter anderen Bedingungen andere Ergebnisse liefern. Des Weiteren ist es möglich, dass Pseudogene mit Ähnlichkeit zu spezifizierenden Proteinen vorhanden sein könnten. Um eine noch umfangreichere Verteilung an spezifizierenden Proteinen aus weiteren Spezies und Familien der Ordnung Brassicales zu erhalten, wurde der DOP-PCR-Ansatz daher trotz fehlender Hinweise auf einfache Nitrile, Epithionitrile oder organische Thiocyanate im phytochemischen Screening der Spezies *A. glabra*, *C. spinosa*, *E. cheiri*, *L. douglasii* und *M. oleifera* durch Templates aus diesen Spezies ergänzt (2.3, 2.10). Die aufgeführten Spezies gehören den Tribus Arabideae, Cardamineae und Erysimeae der Familie der Brassicaceae sowie den Familien Cleomaceae, Limnanthaceae und Moringaceae der Ordnung Brassicales an. Innerhalb des erweiterten DOP-PCR-Ansatzes gelang die Isolierung von zwei *ESP*-Sequenzen aus Spezies der

Familien Cleomaceae und Moringaceae (2.10). Auf Grund ihrer fast 100 % Nukleotidsequenzidentität zu anderen Sequenzen und der darauf begründeten Vermutung, dass diese Sequenzen Artefakte darstellten, wurden sie jedoch nicht in die weiteren phylogenetischen Untersuchungen einbezogen. Weitere Klonierungsversuche blieben ohne Erfolg.

Zur Ergänzung der phylogenetischen Analysen wäre eine Klonierung spezifizierender Proteine aus der Spezies *E. cheiri* allerdings wünschenswert gewesen, da diese Spezies aus dem bisher nicht einbezogenen Tribus Erysimeae der Familie der Brassicaceae stammt. Sehr interessant, nicht nur in Hinsicht auf die Phylogenie der spezifizierenden Proteine, wäre zudem die Klonierung von *ESP*-, *TFP*- oder *NSP*-Sequenzen aus *R. lutea*, *R. luteola* und *T. minus* gewesen, da sich diese Spezies taxonomisch gesehen in die evolutiv sehr alten Familien Resedaceae und Tropaeolaceae der Brassicales eingliedern (Beilstein et al. 2010).

Zusätzlich zu dieser Erweiterung der PCR-Strategie wurde versucht die Ergebnislage durch den durchgeführten Southern-Kapillar-Blot basierend auf genomischen DNA-Fragmenten zu ergänzen (2.10). Mit Hilfe von Sonden für einen spezifischen Nachweis von *ESP*/*TFP*-Sequenzen getrennt von *NSP*-Sequenzen gelang es auch in den Spezies *A. glabra*, *C. microcarpa*, *C. rubella*, *E. sativa*, *E. cheiri*, sowie *R. lutea*, *R. luteola* und *T. minus* das Vorhandensein *ESP*-ähnlicher Sequenzabschnitte nachzuweisen (Abbildung 20, Abbildung 19B). Die Klonierung und Sequenzierung der erhaltenen PCR-Fragmente steht jedoch noch aus. Bestätigt werden konnte zudem die Anwesenheit *ESP*-ähnlicher Sequenzen in Spezies, aus welchen im Vorfeld die Isolierung von *ESPs* oder *ESP*-ähnlichen DNA-Fragmenten gelungen war (*B. vulgaris*, *B. oleracea*, *C. spinosa*, *I. amara*, *I. tinctoria*) (Abbildung 20B). Des Weiteren war der Nachweis von genomischen PCR-Fragmenten mit hoher Ähnlichkeit zu *NSPs* in den Spezies *B. oleracea*, *C. microcarpa*, *E. cheiri*, *E. sativa* und *I. amara* erfolgreich. Die bereits dargestellte Identifizierung eines *NSPs* aus *I. tinctoria* konnte bestätigt werden (Abbildung 20C). Nicht bestätigt werden konnte das Vorkommen einer *ESP*-Sequenz in *M. oleifera* (Abbildung 20B). Diese Tatsache deckt sich mit der Vermutung, dass die Klonierung eines *ESPs* aus *M. oleifera* möglicherweise auf einer nicht näher lokalisierbaren Verunreinigung beruht. Das Ergebnis zeigt deutlich, dass das der Klonierung vorangegangene phytochemische Screening lediglich Hinweise auf das Vorhandensein spezifizierender Proteine liefern konnte. Die Abwesenheit von einfachen Nitrilen, Epithionitrilen oder organischen Thiocyanaten in einem untersuchten Pflanzenorgan ist somit nicht gleichzusetzen mit dem grundsätzlichen Fehlen spezifizierender Proteine in der gesamten Pflanze. Ebenso kann auch durch eine nicht gelungene Klonierung keine Aussage über die Abwesenheit spezifizierender Proteine getroffen werden.

Auf Grund der Tatsache, dass es im zeitlichen Rahmen der vorliegenden Arbeit nicht gelang ausreichend genomische DNA zur direkten Verwendung in einem für den Southern-Kapillar-Blot notwendigen Restriktionsverdau zu isolieren wurde dieser, wie beschrieben, auf Basis von genomischen PCR-Fragmenten als Template durchgeführt. Die Optimierung einer Methode zur Isolierung genomischer DNA für eine Verwendung in einem Southern-Kapillar-Blot ist jedoch unabdingbar, um die Problematik übersehener Hinweise auf spezifizierende Proteine gänzlich auszuschließen. Interessant wäre auch die Identifizierung möglicher Pseudogene mit Ähnlichkeit zu spezifizierenden Proteinen. Ebenfalls interessant

wäre die Generation von Sonden mit einer Spezifität zu *AtNSP5*- und *At3g07720*-ähnlichen Sequenzen für eine Verwendung im Southern-Kapillar-Blot, um ihr Vorhandensein in anderen Arten nachzuweisen. Zudem konnten in einigen Spezies keine spezifizierenden Proteine oder nur jeweils ein ESP, TFP oder NSP oder nur ein cDNA-Fragment identifiziert werden. Die Klonierungen wurden jedoch auf Basis von RNA als Ausgangstemplate vorgenommen. Somit ist es möglich, dass in den untersuchten Pflanzenorganen zum Zeitpunkt der Probenahme eventuelle weitere spezifizierenden Proteine auf Grund des Entwicklungszustandes der Pflanze, des gewählten Organs oder der Wachstumsbedingungen nicht exprimiert waren. Auch in diesem Zusammenhang wäre es daher interessant genomische DNA in die Untersuchungen einzubeziehen. Die Isolierung genomischer DNA würde zudem die Grundlage für die Isolierung genomischer Sequenzen spezifizierender Proteine liefern. An Hand genomischer Sequenzen wäre es möglich, die Intron-Exon-Strukturen der spezifizierenden Proteine zu vergleichen, um weitere Informationen über ihre Evolution zu erhalten. Introns tendieren in höherem Maß zur Akkumulation von Mutationen als Exons, da sie weniger selektiven Einflüssen unterliegen. Somit könnte eine Untersuchung der Introns spezifizierender Proteine weitere Einsichten in ihre Evolution ermöglichen.

Eine genaue Übersicht aller DOP-PCR-Fragmente und der im Southern-Kapillar-Blot erhaltenen Banden für spezifizierende Proteine sowie *Full-length* Sequenzen charakterisierter, der Datenbankrecherche entnommener oder im Vorfeld bekannter spezifizierender Proteine bietet Tabelle 11 unter Angabe von Gattung, Familie und Tribus der untersuchten Spezies sowie im phytochemischen Screening detektierter Glucosinolat-Hydrolyseprodukte. Die im phytochemischen Screening detektierten Glucosinolat-Hydrolyseprodukte der jeweiligen Herkunftspflanze stimmen hierbei mit den Ergebnissen der funktionellen Zuordnung der spezifizierenden Proteine überein. Die einzige Ausnahme bildete das ESP der Spezies *I. tinctoria* in welcher keine Epithionitrile nachweisbar waren.

Tabelle 11: Zusammenfassung der phytochemischen, molekularbiologischen und biochemischen Untersuchungen spezifizierender Proteine. Die untersuchten Arten sind in alphabetischer Reihenfolge mit den jeweiligen experimentellen Ergebnissen aufgeführt. Zur Detektion von Glucosinolat-Hydrolyseprodukte wurden je nach Verfügbarkeit Samen, Blätter, Wurzeln und Keimlinge in wässriger Lösung homogenisiert und Dichlormethanextrakte gewonnener Homogenate mittels GC-MS analysiert. Das Vorhandensein entsprechender Glucosinolat-Hydrolyseprodukte in mindestens einem untersuchten Gewebe ist durch ein + markiert. Die Anzahl erfolgreich charakterisierter sowie im Vorfeld bekannter spezifizierter Proteine der drei Typen ist angegeben. Zusätzlich ist die Anzahl unvollständiger DOP-PCR-Fragmente pro Spezies angegeben. In der Spalte „Blot-Bande“ ist angegeben, mit welcher/n Sonde/n eine Bande im Southern-Kapillar-Blot von PCR-Produkten erhalten wurde. B. = Brassicaceae, ITC = Isothiocyanat, CN = einfaches Nitril, Epithio-CN = Epithionitril, SCN = organisches Thiocyanat, n.t = nicht getestet, n.i = nicht identifizierbar, (*) = vermutlich Artefakt

Spezies	Familie / Tribus	Detektierte Hydrolyseprodukte				Charakterisierte spezifizierende Proteine			Zusätzlich detektiert	
		ITC	CN	Epithio-CN	SCN	NSP	ESP	TFP	DOP-PCR-Fragment	Blot-Bande
<i>Alliaria petiolata</i>	B./Thlaspideae	+	+	+	+			1	1	n.t
<i>Alyssum alpestre</i>	B./Alysseae	+	+						n.t	n.t
<i>Arabis glabra</i>	B./Arabideae	+							n.i	ESP
<i>Barbarea intermediata</i>	B./Cardamineae	+							n.t	n.t
<i>Barbarea stricta</i>		+							n.t	n.t
<i>Barbarea vulgaris</i>		+							1	ESP
<i>Camelina microcarpa</i>	B./Camelineae	n.i.	n.i.	n.i.	n.i.				n.t	ESP / NSP
<i>Capsella bursa-pastoris</i>	B./Camelineae	+							n.t	n.t
<i>Capsella rubella</i>		+							n.t	ESP
<i>Cardamine hirsuta</i>	B./Cardamineae	+	+	+		1	1		1	n.t
<i>Cardamine impatiens</i>		+	+	+			1			n.t
<i>Cleome hassleriana</i>	Cleomaceae	+							n.t	n.t
<i>Cleome spinosa</i>		+					(*)			ESP
<i>Diplotaxis muralis</i>	B./Brassicaceae	+	+						n.i	n.t
<i>Draba aurea</i>	B./Arabideae	+		+			1			n.t
<i>Draba cinerea</i>		+	+						2	n.t

<i>Draba lanceolata</i>		+	+	+		1			n.i	n.t
<i>Draba muralis</i>		+	+						n.i	n.t
<i>Eruca sativa</i>	B./Brassicaceae	+							n.t	ESP / NSP
<i>Erysimum cheiri</i>	B./Erysimeae	+							n.i	ESP / NSP
<i>Erysimum hieraciifolium</i>		+	+						1	n.t
<i>Iberis amara</i>	B./Iberideae	+	+				(*)			ESP / NSP
<i>Isatis tinctoria</i>	B./Isatideae		+			1	1		1	ESP / NSP
<i>Limnanthes douglasii</i>	Limnanthaceae	+							n.i	n.t
<i>Moringa oleifera</i>	Moringaceae	+					(*)			n.i
<i>Reseda lutea</i>	Resedaceae	+							n.t	ESP
<i>Reseda luteola</i>		+							n.t	ESP
<i>Schouwia purpurea</i>	B./Brassicaceae		+	+		1	1			n.t
<i>Thlaspi arvense</i>	B./Thlaspidaceae	+		+	+			1	1	n.t
<i>Tropaeolum</i>	Tropaeolaceae	+	+						n.t	ESP
<u>Ergebnisse der Datenbankrecherche → nicht funktionell charakterisiert</u>										
<i>Arabidopsis lyrata</i>	B./Camelineae	n.t	n.t	n.t	n.t		1		n.t	n.t
<i>Brassica rapa</i>	B./Brassicaceae	n.t	n.t	n.t	n.t		1		n.t	n.t
<i>Eutrema halophilum</i>	B./Eutremeae	n.t	n.t	n.t	n.t	1			n.t	n.t
<u>Im Vorfeld bekannte spezifizierende Proteine</u>										
<i>Arabidopsis thaliana</i>	B./Camelineae	+	+	+		5	1			n.t
<i>Brassica oleracea</i>	B./Brassicaceae	+		+			1			ESP / NSP
<i>Lepidium sativum</i>	B./Lepidieae	+	+		+			1		n.t

3.2 Phylogenetische Analysen spezifizierender Proteine

Durch die Berechnung phylogenetischer Stammbäume unter Verwendung verschiedener Berechnungsgrundlagen war es möglich tiefere Einblicke in die Phylogenie der spezifizierenden Proteine zu erhalten. Hierbei wurden unabhängig vom verwendeten Algorithmus Bäume ähnlicher Topologie erhalten (2.7).

Als Ausgangshypothese für die Analysen wurde die auf Grund erster phylogenetischer Untersuchungen aufgestellte Annahme berücksichtigt, dass die zu AtNSP1 – AtNSP5 (Burow et al. 2009; Kissen and Bones 2009) homologe Sequenz At3g07720 (Burow et al. 2009; Kissen and Bones 2009) ein Vorläuferprotein von AtNSP1 - AtNSP5 darstellt (Burow et al. 2009). Aus diesen Untersuchungen wurde außerdem gefolgert, dass das Vorhandensein der Lektin-ähnlichen Jacalindomänen ein abgeleiteter Zustand ist, da At3g07720 und AtNSP5 diese nicht besitzen (1.5). Eine weitere Annahme von Burow et al. 2009 war, dass die Clusterung der ESPs und TFPs umgeben von AtNSP1 – AtNSP5 sowie At3g07720 einen ersten Hinweis auf die mögliche Erweiterung der Fähigkeit, einfache Nitrile zu synthetisieren, durch die zusätzlich erworbene Möglichkeit zur Formung von Epithionitrilen darstellt (Burow et al. 2009) (1.5, 1.7).

Die Auswahl der Stammbäume erfolgte wie bereits beschrieben entsprechend der besten Bootstrapwerte (Efron 1979) (2.7). Da diese die Verlässlichkeit eines Knotenpunktes innerhalb eines Stammbaumes angeben, kann im Fall der vorliegenden phylogenetischen Untersuchungen für die überwiegende Anzahl der Knotenpunkte eine sichere Darstellung der Verwandtschaftsverhältnisse angenommen werden. Allgemein wird beschrieben, dass bei Bootstrapwerten von 75 % - 80 % und darüber von einer hohen Wahrscheinlichkeit ausgegangen werden kann, dass die Verwandtschaftsverhältnisse stabil widerspiegelt werden. Nicht für jeden Knotenpunkt der berechneten Stammbäume konnten Bootstrapwerte von über 75 % erreicht werden. Da jedoch die mit Hilfe der drei beschriebenen Verfahren auf Basis der Nukleotidsequenzen und auch die auf Basis der Aminosäuresequenzen erstellten Stammbäume eine jeweils vergleichbare Topologie aufweisen (2.7), kann von einer hohen Zuverlässigkeit der vorhergesagten Verwandtschaftsverhältnisse der spezifizierenden Proteine ausgegangen werden.

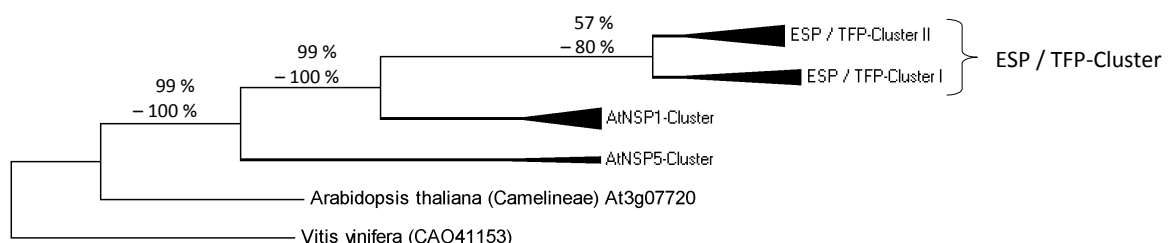


Abbildung 25: Phylogenie spezifizierender Proteine in schematischer Darstellung unter Angabe von Bootstrapwerten. Zur Erstellung verwendet wurden Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der schematisch dargestellte Baum basiert auf den Topologien des Neighbor Joining Baumes, des Maximum Likelihood Baumes und des Maximum Parsimonie Baumes jeweils basierend auf dem Kimura 2-Parameter-Modell. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Die Darstellung zeigt das grundlegende Clusterverhalten der Sequenzen sowie die Spanne der ermittelten Bootstrapwerte aus allen drei verwendeten Berechnungsmodellen mit 1000facher Wiederholung. Astlängen geben die geschätzten evolutiven Distanzen wieder.

Durch die erfolgreiche Erweiterung des gegebenen Datensatzes aus acht spezifizierenden Proteinen (Tabelle 1) auf insgesamt 23 Nukleotid- und Aminosäuresequenzen spezifizierender Proteine bekannter Herkunft und Funktion (eine Ausnahme bilden die drei der Datenbankrecherche entnommenen Sequenzen und At3g07720) (Tabelle 6, Tabelle 7), ergibt sich mit hoher Wahrscheinlichkeit die bereits im Detail beschriebene Clusterung der *ESP*- mit den *TFP*-Sequenzen (*ESP* / *TFP*-Cluster) vollständig getrennt von den Sequenzen des *NSP*-Clusters (2.7). An der Basis steht At3g07720 (Abbildung 25). Hierbei ist At3g07720 jedoch nicht als alleinstehende Sequenz zu verstehen sondern als Vertreter einer Gruppe ähnlicher jedoch im Rahmen der vorliegenden Arbeit nicht genauer untersuchter Homologe, die auch in nicht-Glucosinolat-haltigen Pflanzen vorkommen. Auf Grund der in größerer Anzahl vorliegenden *NSP*-Sequenzen wird eine Weiterentwicklung innerhalb der *NSPs* ausgehend von AtNSP5 und homologen Sequenzen (AtNSP5-Cluster) durch die Aufspaltung in die weiteren *NSPs* (AtNSP1-Cluster) deutlich (Abbildung 25). Das abstammungsgeschichtlich ältere AtNSP5-Cluster zeigt hier an Hand seiner Astlänge eine deutliche evolutive Distanz zu At3g07720 und dessen Homologen. Die Sequenzen entstammen den Tribus Camelineae (AtNSP5) und Eutremeae (EhNSP) (Taji et al. 2008) der Abstammungslinien I und II (Al-Shehbaz et al. 2006; Beilstein et al. 2006). Die Abzweigung zwischen At3g07720 und dem AtNSP5-Cluster ist mit einem Bootstrapwert von 98 %- 100 % in allen berechneten Topologien anzugeben (Abbildung 25). Eine ähnliche evolutive Distanz zu At3g07720 und seinen Homologen weist auch das mit Bootstrapwerten von 99 % - 100 % vom AtNSP5-Cluster abgespaltene AtNSP1-Cluster auf (Abbildung 25). Das AtNSP1-Cluster enthält Vertreter der Tribus Camelineae, Cardamineae, Isatideae und Brassiceae aus den Abstammungslinien I und II (Al-Shehbaz et al. 2006; Beilstein et al. 2006). Das aus dem AtNSP1-Cluster hervorgehende *ESP* / *TFP*-Cluster hingegen zeigt mit Bootstrapwerten von 100 % eine größere evolutive Distanz zu At3g07720 und seinen Homologen, erkennbar an seiner längeren Astlänge (Abbildung 25). Interessanterweise ließen sich in der Phylogenie durch den erweiterten Datensatz der spezifizierenden Proteine nicht nur die *NSP*-Sequenzen in die zwei Gruppen des AtNSP1- und des AtNSP5-Clusters unterteilen sondern auch das *ESP* / *TFP*-Cluster zeigte eine Aufspaltung in die Cluster *ESP* / *TFP*-Cluster I und *ESP* / *TFP*-Cluster II (Abbildung 25). Das *ESP* / *TFP*-Cluster II beinhaltet spezifizierende Proteine der Tribus Arabideae, Brassiceae, Isatideae, Thlaspidae der Abstammungslinie II (Al-Shehbaz et al. 2006; Beilstein et al. 2006) und das *ESP* / *TFP*-Cluster I Proteine der Tribus Camelineae, Cardamineae, Lepidieae der Abstammungslinie I (Al-Shehbaz et al. 2006; Beilstein et al. 2006). Beide Untercluster zeigen ähnliche Astlängen (Abbildung 25). Die innerhalb des *ESP* / *TFP*-Clusters enthaltenen *TFP*-Sequenzen stehen sich dabei über beide Untercluster als Paraphylum gegenüber (Abbildung 11, Abbildung 12, Abbildung 13). Die Abzweigung der beiden Untercluster weist jedoch nur Bootstrapwerte von 57 % - 80 % an ihrem Knotenpunkt auf (Abbildung 25). Wegen ihres Auftretens in allen berechneten Baumtopologien erscheint aber auch diese Aufspaltung sicher.

Bei der Betrachtung der phylogenetischen Bäume und der im Folgenden diskutierten Distanzdaten ist zu beachten, dass Lücken in Alignments vollständig vernachlässigt wurden (Maximum Parsimonie-Methode) oder vernachlässigt wurden, wenn mehr als 95 % der Sequenzen in einer Position uninformativ waren (Neighbor Joining-Verfahren, Maximum Likelihood-Methode). Die dargestellten Stammbäume (Abbildung 11 bis 13 sowie Abbildung

15) repräsentieren daher im Wesentlichen die Phylogenie der Kelch-Domänen. Belegt wurde diese Tatsache durch die Berechnung eines Stammbaumes unter nukleotidgenauem Ausschluss der Lektin-ähnlichen Jacalindomänen, welcher eine identische Baumtopologie zeigte (Abbildung 16). Zum Auftreten der beschriebenen Clusterung spezifizierender Proteine kommt es also unabhängig vom Vorhandensein der Lektin-ähnlichen Jacalindomänen bei den NSPs des AtNSP1-Clusters (Zur Phylogenie der Lektin-ähnlichen Jacalindomänen siehe Abbildung 17 sowie Abschnitt 3.2.2).

Die beschriebene Anordnung der einzelnen Cluster innerhalb der Baumtopologie bestätigte die Annahme, dass die Fähigkeit Epithionitrile zu bilden ein von den NSPs abgeleitetes funktionserweiterndes Merkmal der ESPs darstellt. Auf Basis der ESPs könnte sich dann die Fähigkeit organische Thiocyanate zu formen entwickelt und zu TFP-Aktivität geführt haben.

Das sortierte Vorkommen von ESPs in den zwei Unterclustern des ESP / TFP-Clusters entsprechend den Abstammungslinien I und II der Brassicaceae zeigt zudem, dass eine Genduplikation, welche zu dem Split zwischen dem AtNSP1-Cluster und dem ESP / TFP-Cluster führte, bereits vor der Auffächerung der Kerngruppe der Familie Brassicaceae in die Abstammungslinien I und II stattgefunden hat. Auf Grund der Tatsache, dass die Differenzierung der Kerngruppe der Familie Brassicaceae in die beiden Abstammungslinien auf ~ 35,6 Mill. Jahre (I) und ~ 30,8 Mill. Jahre (II) eingeordnet wurde (Beilstein et al. 2010), kann das Alter der Differenzierung der NSPs des AtNSP1-Clusters und der ESPs des ESP / TFP- Clusters auf mindestens 35,6 Mill. Jahre geschätzt werden. Der Split, welcher ausgehend von At3g07720 als Vertreter einer Gruppe ähnlicher Homologe zu den ersten im Glucosinolat-Myrosinase-System funktionsfähigen NSPs des AtNSP5-Clusters führte sowie der Split, welcher ausgehend von den NSPs des AtNSP5-Clusters zur Entstehung der NSPs des AtNSP1-Clustes führte, müssen entsprechend älter sein.

3.2.1 Ist At3g07720 ein Vorläuferprotein spezifizierender Proteine?

Durch die Bewurzelung der generierten Stammbäume mit der Sequenz CAO41153 aus *V. vinifera* (Jaillon et al. 2007), konnte die Stellung von At3g07720 (Burow et al. 2009; Kissen and Bones 2009) als Homolog der NSPs bestätigt werden (Abbildung 26).

Untermuert wird diese Aussage durch die eher geringen Sequenzidentitäten dieses Proteins mit den NSPs, ESPs und TFPs des Glucosinolat-Myrosinase-Systems (Aminosäureebene: 44 % - 57 %, Nukleotidebene: 61 % - 67 %) (Tabelle S 2). In der Umkehrung wird dies auch durch die ermittelten Sequenzdifferenzen als Maß evolutiver Änderungen zwischen zwei Sequenzen bestätigt, welche für diese Vergleiche hohe Werte lieferten (Aminosäureebene: 82,7 % - 146,1 % Nukleotidebene: 78,3 % – 123,0 %) (Tabelle 8, Tabelle S 3). Außerdem kommen neben dem Protein CAO41153 aus *V. vinifera* weitere homologe Sequenzen mit 62 % Identität auf Nukleotidebene (Aminosäureebene: 60 % - 63 %) zu At3g07720 in Gattungen außerhalb der Brassicales vor, die keine Glucosinolate enthalten (vergleiche auch Abbildung 6), z.B. in *Hordeum vulgare* (Poaceae) und *Oryza sativa* (Poaceae).

Die vorliegenden phylogenetischen Analysen bestätigen somit die Vermutung, dass At3g07720 der Repräsentant einer Gruppe von Vorläuferproteinen ist, welche die gemeinsamen Vorfahren aller spezifizierenden Proteine darstellen jedoch keine Funktion bei der Glucosinolat-Hydrolyse sondern eine andere, bisher nicht identifizierte Funktion haben. Bisher wurde keines dieser Proteine charakterisiert. Anhaltspunkte für eine Funktion von At3g07720 beziehen sich bisher ausschließlich auf ein erhöhtes Transkriptionslevel in Wurzeln von *A. thaliana* bei Eisenmangel (Colangelo and Guerinot 2004; Buckhout et al. 2009; Stein and Waters 2011). Im Rahmen der vorliegenden Arbeit wurde nicht gezielt versucht weitere Sequenzen mit hoher Ähnlichkeit zu At3g07720 aus weiteren Spezies der Ordnung Brassicales oder anderer Ordnungen zu identifizieren.

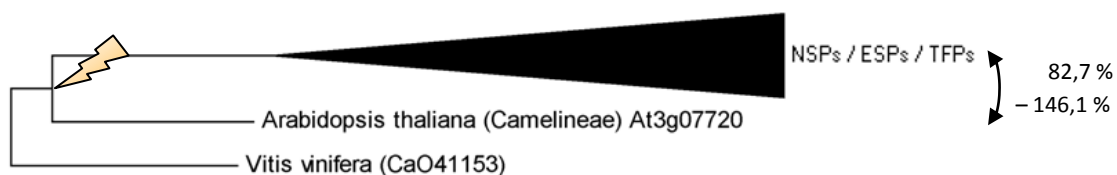


Abbildung 26: Phylogenie spezifizierender Proteine und ihrer Vorläufer in schematischer Darstellung. Zur Erstellung verwendet wurden die Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der schematisch dargestellte Baum basiert auf den Topologien des Neighbor Joining Baumes, des Maximum Likelihood Baumes und des Maximum Parsimonie Baumes jeweils basierend auf dem Kimura 2-Parameter-Modell. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Die Darstellung zeigt das grundlegende Clusterverhalten der Sequenzen. Astlängen geben die geschätzten evolutiven Distanzen wieder. Der Split von At3g07720 zu NSP-, ESP- und TFP-Sequenzen ist mit einem Pfeil markiert (orange). Rechts ist die Sequenzdivergenz auf Aminosäureebene angegeben.

3.2.2 Sind NSPs die Basis der Entwicklung zu ESPs und TFPs?

Dargestellt wurde bereits, dass At3g07720 (Burow et al. 2009; Kissen and Bones 2009) und Homologe an der Basis der phylogenetischen Entwicklung der spezifizierenden Proteine stehen und sich aus ihnen NSPs entwickelt haben. Die phylogenetische Analyse zeigte des Weiteren, dass sich NSPs in die zwei Gruppen des AtNSP5-Clusters und des AtNSP1-Clusters unterteilen lassen (Abbildung 27).

Diese Unterteilung wird unterstrichen durch die berechneten Werte für die Sequenzidentitäten (2.6.1) sowie die Sequenzdivergenzen (2.8) innerhalb der beiden Cluster in Gegenüberstellung zu einem Vergleich zwischen den beiden Clustern. Hier zeigen sich innerhalb des AtNSP5-Clusters (Identität: Aminosäureebene: 85 %, Nukleotidebene: 84 %; Divergenz: Aminosäureebene: 14,9 %, Nukleotidebene: 21,5 %) (Tabelle 8, Tabelle S 2, Tabelle S 3) und innerhalb des AtNSP1-Clusters (Identität: Aminosäureebene: 74 % - 92 %, Nukleotidebene: 82 % - 95 %; Divergenz: Aminosäureebene: 5,7 % - 28,6 %, Nukleotidebene: 4,1 % - 21,2 %) (Tabelle 8, Tabelle S 2, Tabelle S 3) Werte, welche eine enge Verwandtschaft der spezifizierenden Proteine je eines Clusters widerspiegeln, während der Vergleich der NSPs verschiedener Cluster deutlich auf eine größere evolutive Distanz hinweist (Identität: Aminosäureebene: 50 % - 54 %, Nukleotidebene: 66 % - 70 %; Divergenz: Aminosäureebene: 95,5 % - 107,4 %, Nukleotidebene: 79,4 % - 92 %) (Tabelle 8, Tabelle S 2, Tabelle S 3).

Auch die vorhergesagte Domänenstruktur (2.6.2) aus vier Kelchdomänen in allen Sequenzen des AtNSP1-Clusters und nur drei Kelchdomänen für At3g07720 und die Sequenzen des AtNSP5-Clusters stimmt mit dieser Aussage überein (Abbildung 4, Abbildung 10). Außerdem sind in allen Sequenzen des AtNSP1-Clusters Lektin-ähnliche Jacalindomänen vorhanden, während diese bei den Sequenzen des AtNSP5-Clusters und At3g07720 fehlen (Abbildung 4, Abbildung 10). Dies bestätigt die Hypothese, dass das Vorhandensein der Lektin-ähnlichen Jacalindomänen ein abgeleiteter Zustand ist (Burow et al. 2009), da das AtNSP5-Cluster das abstammungsgeschichtlich ältere Cluster ist. Innerhalb der NSPs scheint somit eine von den NSPs des AtNSP5-Clusters weiterführende strukturelle Entwicklung zu den NSPs des AtNSP1-Clusters stattgefunden zu haben. Die Produktspezifität der weiterentwickelten NSPs ist jedoch nach heutigem Wissensstand identisch mit derjenigen der ursprünglichen NSPs (Burow et al. 2009) (2.4).

Für Erkenntnisse über die evolutive Abstammung der Lektin-ähnlichen Jacalindomänen wurde in Anlehnung an Untersuchungen von Burow et al. 2009 (Burow et al. 2009) ein phylogenetischer Stammbaum erstellt, welcher ausschließlich die Lektin-ähnlichen Jacalindomänen betrachtet (Abbildung 17). Wie bereits in Burow et al. 2009 wurde die N-terminal gelegene Lektin-ähnlichen Jacalindomänen des AtNSP4 mit AtNSP4A, die darauf folgende mit AtNSP4B bezeichnet. Bestätigt werden konnte, dass AtNSP4A einer unbekannten Quelle entstammt, während AtNSP4B und die Lektin-ähnlichen Jacalindomänen aus AtNSP1 – AtNSP3 ihren Ursprung in putativen MBPs aus *A. thaliana* haben (Burow et al. 2009). Die Lektin-ähnlichen Jacalindomänen aller neu identifizierten NSPs sortieren sich ebenfalls den putativen MPBs zu (Abbildung 17) und haben somit einen gemeinsamen Ursprung mit der Lektin-ähnlichen Jacalindomäne B des AtNSP4 und den

Lektin-ähnlichen Jacalindomänen aus AtNSP1 – AtNSP3. Dies unterstreicht eine Weiterentwicklung des AtNSP4 durch den Zugewinn der zweiten Lektin-ähnlichen Jacalindomäne. Weiterführend wurde eine Distanzberechnung für die Lektin-ähnlichen Jacalindomänen vorgenommen (2.8). Die erhaltenen Zahlenwerte bestätigen die Baumtopologie und zeigen die bereits beschriebene Herkunft der zwei Lektin-ähnlichen Jacalindomänen des AtNSP4 aus zwei unterschiedlichen Quellen (Burow et al. 2009). Der Vergleich der Lektin-ähnlichen Jacalindomäne B des AtNSP4 mit den Lektin-ähnlichen Jacalindomänen der NSPs des AtNSP1-Clusters (ohne AtNSP4B) zeigte Distanzen zwischen 11,9 % und 71,0 % auf Aminosäureebene (Nukleotide: 8,3 % - 34,8 %) (Tabelle S 5). Vergleicht man hingegen den Sequenzabschnitt für die Lektin-ähnliche Jacalindomäne A des AtNSP4 mit den Lektin-ähnlichen Jacalindomänen der NSPs des AtNSP1-Clusters (ohne AtNSP4A), so wird an Hand der erhöhten Distanzwerte und somit zunehmenden Aminosäure- und Nukleotidaustauschen die geringere Verwandtschaft deutlich (Aminosäuren: 164,0 % - 249,7 %, Nukleotide: 157,9 % - 201,1 %) (Tabelle S 5).

An Hand der berechneten Baumtopologie zeigt sich des Weiteren, dass nur die Sequenzen des AtNSP1-Clusters direkte Vorläufer der ESPs und TFPs sind (Abbildung 27). Somit ergibt sich eine Aufspaltung ausgehend von dem AtNSP5-Cluster in das AtNSP1-Cluster und ausgehend von diesem in das ESP / TFP-Cluster, wobei nur zwei Vertreter für das AtNSP5-Cluster für diese Analyse zur Verfügung standen (Abbildung 27).

Belegen lässt sich diese Aussage wiederum durch die Vergleiche der Sequenzidentitäten (2.6.1) und der Sequenzdivergenzen (2.8) zwischen den NSPs des AtNSP5-Clusters jeweils mit denjenigen des AtNSP1-Clusters und den *ESP*-Sequenzen des ESP / TFP-Clusters. Aus diesen Berechnungen ergibt sich auf Ebene der Aminosäuren eine höhere Ähnlichkeit der ESPs des ESP / TFP-Clusters zu den NSPs des AtNSP1-Clusters (Identität: 50 % - 60 %, Divergenz: 60,5 % - 98,2 %) (Tabelle 8, Tabelle S 2, Tabelle S 3) als zu den NSPs des AtNSP5-Clusters (Identität: 46 % - 51 %, Divergenz: 107,3 % - 135,9 %) (Tabelle 8, Tabelle S 2, Tabelle S 3). Die NSPs des NSP5-Clusters sind zudem enger mit den *NSP*-Sequenzen des AtNSP1-Clusters verwandt (Identität: 50 % - 54 %; Divergenz: 95,5 % - 107,4 %) (Tabelle 8, Tabelle S 2, Tabelle S 3) als mit den *ESP*-Sequenzen. Die gleiche Tendenz ließ sich auch auf Ebene der Nukleotide nachweisen, jedoch nicht so deutlich ausgeprägt (Tabelle S 2, Tabelle S 3).

Auch die vorhergesagte Domänenstruktur (2.6.2) der ESPs, TFPs und NSPs des AtNSP1-Clusters hat sich, im Vergleich zu At3g07720 und den NSPs des AtNSP5-Clusters verändert (Abbildung 4, Abbildung 10). Diese Tatsache kann aus der vorhergesagten Ausprägung von drei Kelch-Domänen in At3g07720, AtNSP5 und EhNSP und von vier Kelchdomänen in ESPs, TFPs und den NSPs des AtNSP1-Clusters abgeleitet werden (Abbildung 4, Abbildung 10). Somit weisen ESPs und TFPs in ihren strukturellen Elementen große Ähnlichkeiten zu den *NSP*-Sequenzen des AtNSP1-Clusters auf, besitzen jedoch keine Lektin-ähnlichen Jacalindomänen (Abbildung 4, Abbildung 10). Da alle *NSP*-Sequenzen des AtNSP1-Clusters Lektin-ähnliche Jacalindomänen aufweisen scheinen sie früh in der evolutiven Entwicklung des AtNSP1-Clusters aufgetreten zu sein. Auf Grund ihres Fehlens in den *ESP*- und *TFP*-Sequenzen entstanden Lektin-ähnlichen Jacalindomänen jedoch vermutlich erst nach der Aufspaltung des AtNSP1-Clusters und des

ESP / TFP-Clusters und somit vor weniger als 36 Mill. Jahren. Die strukturellen Veränderungen beim Übergang der Sequenzen des NSP1-Clusters zu ESPs sind weniger ausgeprägt als beim Übergang der NSPs des NSP5-Clusters zu NSPs des NSP1-Clusters, jedoch ist die Produktspezifität der ESPs und TFPs deutlich weiterentwickelt. ESPs und TFPs unterlagen nach der Abspaltung aus NSPs des AtNSP1-Clusters einer Subneofunktionalisierung, bei welcher die Funktion einfache Nitrile zu formen um die Möglichkeit zur Bildung von Epithionitrilen (ESP) oder Epithionitrilen und organischen Thiocyanaten (TFP) erweitert wurde (2.4).

Die NSPs des AtNSP5-Clusters und homologe Sequenzen sind somit Vorläufer der strukturell veränderten Gruppe von Sequenzen des AtNSP1-Clusters mit nach heutigem Wissensstand gleicher biochemischer Funktion. Des Weiteren sind die Sequenzen des AtNSP1-Clusters die Vorläufer der ESPs und TFPs und bildeten die Grundlage für eine Subneofunktionalisierung dieser Proteine. Die untersuchten ESPs, TFPs und NSPs schließen die Tribus Eutremeae, Camelineae, Cardamineae, Brassiceae, Isatideae, Arabideae, Lepidieae und Thlaspeidae der Ordnung Brassicales und somit drei Tribus der Abstammungslinie I (Camelineae, Cardamineae und Lepidieae) sowie sechs Tribus der Abstammungslinie II (inklusive ihrer Erweiterung (Arabideae, Brassiceae, Eutremeae, Isatideae und Thlaspeidae) ein (Al-Shehbaz et al. 2006; Beilstein et al. 2006).

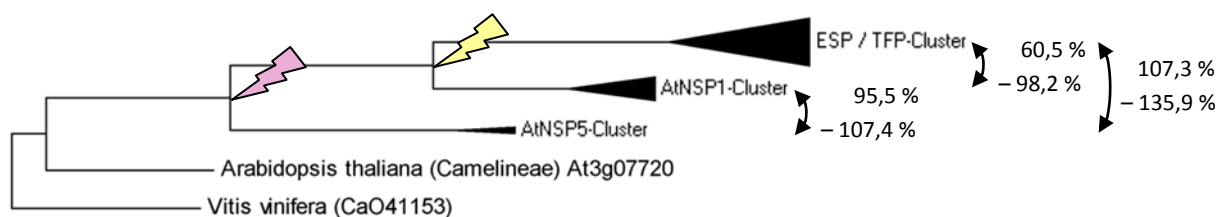


Abbildung 27: Schematische Phylogenie spezifizierender Proteine. Zur Erstellung verwendet wurden die Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der schematisch dargestellte Baum basiert auf den Topologien des Neighbor Joining Baumes, des Maximum Likelihood Baumes und des Maximum Parsimonie Baumes jeweils basierend auf dem Kimura 2-Parameter-Modell. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Die Darstellung zeigt das grundlegende Clusterverhalten der Sequenzen. Astlängen geben die geschätzten evolutiven Distanzen wieder. Der jeweilige Split ausgehend von NSPs des AtNSP5-Clusters zu NSP-Sequenzen des AtNSP1-Clusters (lila) und von den NSPs des AtNSP1-Clusters zu dem ESP / TFP-Cluster (gelb) ist mit einem Pfeil markiert. Rechts ist die Sequenzdivergenz auf Aminosäureebene angegeben.

3.2.3 Sind TFPs zwei Mal unabhängig voneinander aus ESPs entstanden oder kam es zu einem multiplen Verlust?

Bei genauerer Betrachtung des ESP / TFP-Clusters kommt es zu einer Unterteilbarkeit in zwei Untercluster, in welchen sich LsTFP (Burow et al. 2007a) mit ApTFP und TaTFP (Kuchernig et al. 2011) als Paraphylum gegenübersteht (Abbildung 28). Es ergibt sich somit keine Auftrennung in ein ESP-enthaltendes Cluster und ein TFP-enthaltendes Cluster sondern eine Auflösung des Stammbaumes nach Tribus. Dabei gruppieren sich jeweils die spezifizierenden Proteine der Abstammungslinie I (Camelineae, Cardamineae, Lepidieae) und die spezifizierenden Proteine der Abstammungslinie II (Brassiceae, Isatideae, Thlaspidiae) (Beilstein et al. 2006; Beilstein, Al-Shehbaz et al. 2006) (Abbildung 28).

Die berechneten Sequenzdivergenzen (2.8) zwischen NSPs und ESPs (Aminosäureebene: 60,5 % - 135,9 %, Nukleotidebene: 63,7 % - 115,3 %) (Tabelle 8, Tabelle S 3) im Vergleich zu Divergenzen zwischen NSPs und TFPs (Aminosäureebene: 79,0 % - 150,8 %, Nukleotidebene: 70,7 % - 123,8 %) (Tabelle 8, Tabelle S 3) zeigen, dass der Verwandtschaftsgrad zwischen NSPs und TFPs geringer ist als zwischen NSPs und ESPs. TFPs scheinen jedoch eng verwandt mit ESPs zu sein (Aminosäureebene: 28,2 % - 60,5 %, Nukleotidebene: 19,6 % - 44,5 %) (Tabelle 8, Tabelle S 3). Somit stellen TFPs die jüngste Entwicklungsstufe der spezifizierenden Proteine ausgehend von ESPs dar.

Berechnungen der Sequenzidentitäten (2.6.1) und Sequenzdivergenzen (2.8) zwischen ESPs und TFPs zeigen zudem, dass TFPs zwei Mal unabhängig voneinander aus ESPs hervorgegangen sind. Diese Aussage lässt sich mit den erhaltenen Zahlenwerten begründen. ESPs innerhalb des ESP / TFP-Clusters II ebenso wie innerhalb des ESP / TFP-Clusters I und auch über beide Untercluster verglichen (Identität: Aminosäureebene: 71 % - 98 %, Nukleotidebene: 78 % - 98 %; Divergenz: Aminosäureebene: 1,0 % - 41,4 %, Nukleotidebene: 0,8 % - 31,9 %) (Tabelle 8, Tabelle S 2, Tabelle S 3) stehen in engerer Verwandtschaft beziehungsweise kleinerer Distanz zueinander, als im Vergleich mit den TFPs innerhalb der jeweiligen Abstammungslinie als auch über die Abstammungslinien hinweg (Identität: Aminosäureebene: 63 % - 76 %, Nukleotidebene: 74 % - 83 %; Divergenz: Aminosäureebene: 28,2 % - 60,5 %, Nukleotidebene: 19,6 % - 44,5 %) (Tabelle 8, Tabelle S 2, Tabelle S 3). Dies mag unter Umständen darin begründet sein, dass ESPs über eine identische Produktspezifität (Bildung von einfachen Nitrilen und Epithionitrilen) verfügen (Lambrix et al. 2001; Matusheski et al. 2006) (2.4), wohingegen TFPs zusätzlich organische Thiocyanate formen (Burow et al. 2007a; Kuchernig et al. 2011) (2.4). Jedoch weisen die TFPs verschiedener Abstammungslinien gegeneinander verglichen eine größere Distanz und somit einen geringeren Verwandtschaftsgrad auf (Identität: Aminosäureebene: 59 % - 60 %, Nukleotidebene: 73%; Divergenz: Aminosäureebene: 68,1 % und 70,1 %, Nukleotidebene: 53,1 % und 54,5 %) (Tabelle 8, Tabelle S 2, Tabelle S 3) als ESPs und TFPs innerhalb einer Abstammungslinie.

Dies steht in Übereinstimmung mit der zweifach unabhängigen Entstehung von TFPs aus ESPs in der Abstammungslinie I (Lepidieae) und der Abstammungslinie II (Thlaspidiae) (Al-

Shehbaz et al. 2006; Beilstein et al. 2006). Damit lässt sich ihre Entwicklung auf die Zeit nach Auffächerung der Kerngruppe der Familie Brassicaceae in die Abstammungslinien I und II datieren.

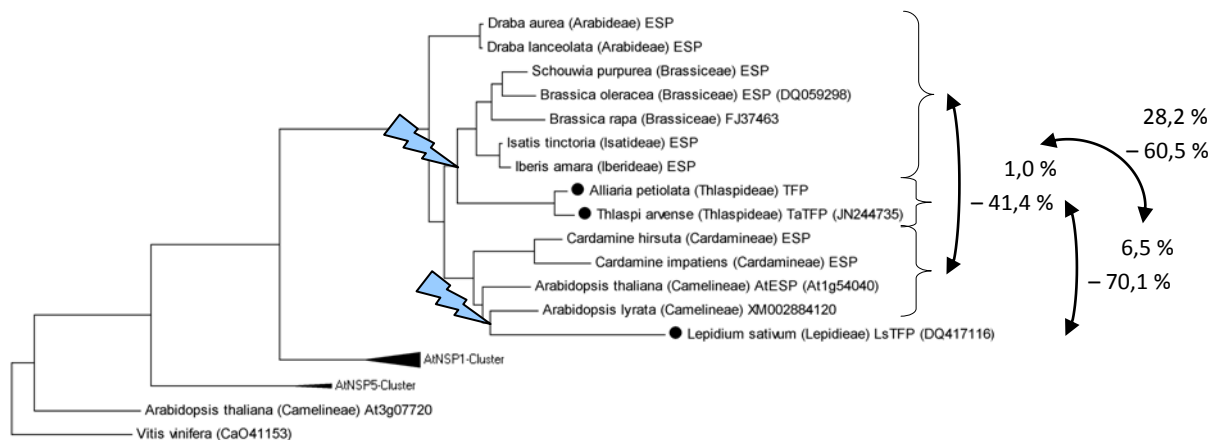


Abbildung 28: Phylogenie spezifizierender Proteine. Zur Erstellung verwendet wurden die Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales von Start- bis Stoppcodon. Der schematisch dargestellte Baum basiert auf den Topologien des Neighbor Joining Baumes, des Maximum Likelihood Baumes und des Maximum Parsimonie Baumes jeweils basierend auf dem Kimura 2-Parameter-Modell. CAO41153 aus *V. vinifera* diene als Außengruppe zur Bewurzelung. Die Darstellung zeigt das grundlegende Clusterverhalten der Sequenzen. Astlängen geben die geschätzten evolutionen Distanzen wieder. Jeder Split ausgehend von ESPs zu TFPs ist mit einem Pfeil markiert (blau). Die betrachteten TFPs sind durch schwarze Kreise hervorgehoben. Rechts ist die Sequenzdivergenz auf Aminosäureebene angegeben.

3.3 Mechanismen der Evolution der spezifizierenden Proteine

Die mutierte Variante eines Proteins unterliegt negativer Selektion, wenn nichtsynonyme Substitutionen sich nachteilig auf den Phänotyp eines Organismus auswirken. Negative Selektion dient der Bereinigung nicht-funktioneller oder unvorteilhafter Proteine. Die Aminosäureabfolge innerhalb der funktionell wichtigen Bereiche eines Proteins ist daher meistens sehr invariabel, da negative Selektion dafür sorgt, dass die Funktion des Proteins erhalten wird. Es treten somit gegenüber synonymen Substitutionen pro synonyme Sequenzposition (d_S) nur wenige bis keine nichtsynonymen Substitutionen pro nichtsynonymer Sequenzposition (d_N) auf. Nachzuweisen ist diese Tatsache durch einen ω -Wert ($\omega = d_N / d_S < 1$) (Yang et al. 2000).

Aus dem vorliegende Datensatz ergaben sich stets Werte von $\omega < 1$ ($\omega = 0,123 - 0,627$) (Tabelle S 6). Durch die hohe Homogenität dieser Werte können sie kaum unterstützend für weiterführende Interpretationen angeführt werden. Auf Grund der Tatsache, dass die ω -Werte (Tabelle S 6) für den Vergleich der NSPs untereinander ($\omega = 0,123 - 0,435$), der NSPs verglichen mit den ESPs ($\omega = 0,219 - 0,351$) und der NSPs verglichen mit TFPs ($\omega = 0,234 - 0,346$) kleinere Werte lieferten als die Vergleiche der ESPs untereinander ($\omega = 0,143 - 0,627$), der TFPs untereinander ($\omega = 0,198 - 0,524$) und der ESPs mit den TFPs ($\omega = 0,344 - 0,549$) kann jedoch gemutmaßt werden, dass auf die *NSP*-Sequenzen eventuell eine stärkere negative Selektion wirkt als auf *ESP*- und *TFP*-Sequenzen. Die Ursache hierfür könnte in der reduzierten funktionellen Einschränkung der ESPs und TFPs durch den Zugewinn neuer biochemischer Funktionen liegen. Um eine tatsächliche Aussage über die von negativer Selektion betroffenen Aminosäurereste der spezifizierenden Proteine zu erhalten wäre die Berechnung der ω -Werte in einer sogenannten „Sliding-Window-Analyse“ notwendig. Eine solche Berechnung dient der lokalen Analyse von ω -Werten über die Länge einer Sequenz.

Auf Grund der Tatsache, dass alle im Rahmen der vorliegenden Arbeit berechneten ω -Werte < 1 (Tabelle S 6) sind wird jedoch deutlich, dass alle Gene der spezifizierenden Proteine des Glucosinolat-Myrosinase-Systems negativer Selektion unterliegen und daher eine positive Wirkung für die Pflanze haben, welche an die Aminosäuresequenzen der Gene gebunden ist. Pflanzen erreichen durch die Vielzahl sowohl der Glucosinolate als auch ihrer Abbauprodukte eine Spezies- sowie Standort- und Entwicklungsstand-spezifische Mischung an Glucosinolaten und Glucosinolat-Hydrolyseprodukten. Diese wird zudem von biotischen und abiotischen Faktoren beeinflusst (Brown et al. 2003; Charron et al. 2005; Charron et al. 2005; Himanen et al. 2008; Wentzell and Kliebenstein 2008; Textor and Gershenzon 2009). Durch die Umlenkung der Glucosinolat-Myrosinase-Reaktion von Isothiocyanaten zu einfachen Nitrilen, Epithionitrilen und organischen Thiocyanaten ermöglichen die spezifizierenden Proteine eine größere Variabilität der entstehenden Glucosinolat-Hydrolyseprodukte. Die Diversität dieser sekundären Metabolite scheint notwendig um auf die die Pflanze umgebenden und wechselnden Umwelteinflüsse entsprechend reagieren zu können (Kroymann 2011). Ihr vielfältiges Auftreten erscheint möglicherweise im evolutiven Wettrennen als aktuell sinnvollste Lösung beim Ausbalancieren der Verteidigung der Pflanzen gegen Schädlinge durch direkte und indirekte

Abwehr. Eine endgültige Klärung der Frage, warum Pflanzen in die Synthese der unterschiedlichen spezifizierenden Proteine investieren steht noch aus. Da jedoch NSP- und ESP-Aktivität zusammen in einer Pflanze nachgewiesen werden konnten, erscheinen die unterschiedlichen biologischen Rollen von einfachen Nitrilen und Epithionitrilen sicher zu sein. Als weitere mögliche Antwort bleibt zudem das Konzept der Coevolution bestehen (Ehrlich and Raven 1964).

In der phylogenetischen Entwicklung stellt die Verdopplung eines Gens eine Möglichkeit dar, negativer Selektion zu entkommen. Ausgelöst werden kann dies von Mechanismen wie der Verdopplung einzelner Gene (Genduplikation), von Chromosomensegmenten, ganzen Chromosomen (Chromosomenduplikation) oder sogar ganzen Genomen (Genomduplikation). Ein Genduplikat unterliegt in jedem Fall einem wesentlich geringeren Selektionsdruck, da es zusätzlich vorhanden ist. Durch die Verdopplung eines Gens bietet sich innerhalb einer Genkopie die Möglichkeit zur Akkumulation von Mutationen, welche neue Funktionen ergeben können, während in der anderen Genkopie die alte Funktion erhalten bleibt. Die Verdopplung eines Gens kann dabei sowohl zu Pseudogenisierungen (Funktionsverlust), Funktionsverstärkungen (Vorliegen von zwei Genen gleicher Funktionalität) als auch zu Varianten neuer Funktionalität führen. Möglichkeiten für die Ausprägung neuer Funktionen liegen dabei in der Neofunktionalisierung (eine Genkopie behält die alte Funktion bei, die zweite Genkopie erlangt eine neue Funktion), Subfunktionalisierung (beide Genkopien teilen sich komplementär die alte Funktion) oder Subneofunktionalisierung (beide Genkopien teilen sich komplementär die alte Funktion unter zusätzlicher Erlangung einer neuen Funktionalität).

Für die im Rahmen der vorliegenden Arbeit untersuchten und im Vorfeld bekannten NSPs, ESPs und TFPs sind die Mechanismen, welche zu der Ausprägung neuer Funktionen der ESPs und TFPs geführt haben, nur schwer zu benennen. Eine Genduplikation scheint zu der Aufspaltung der NSPs des AtNSP5-Clusters zu NSPs des AtNSP1-Clusters geführt zu haben. Diese führte jedoch nicht zu einer funktionellen Erweiterung, sondern zog ausschließlich strukturelle Veränderungen nach sich. Im Fall dieser Genduplikation ist von einer Funktionsverstärkung auszugehen. Eine zweite Genduplikation ergab die Aufspaltung von NSPs des AtNSP1-Clusters zu den ESPs des ESP / TFP-Clusters und ermöglichte die zusätzliche Fähigkeit, Epithionitrile zu formen. ESPs haben dabei jedoch die ursprüngliche Funktion der NSPs, einfache Nitrile zu formen, nicht verloren. Daher scheint in diesem Fall eine Subneofunktionalisierung vorzuliegen. Auch während der Entwicklung der TFPs aus ESPs wurde erneut eine Funktion (die Formung organischer Thiocyanate) zu den vorhergehenden Funktionen der ESPs addiert. Über den Mechanismus ihrer Entstehung lässt sich jedoch auf Grund des Fehlens von paralogenen ESP / TFP-Paaren keine Aussage treffen.

Die Diversifizierung der NSPs des AtNSP1-Clusters und die Entstehung der ESPs des ESP / TFP-Clusters konnte auf die Zeit vor mindestens 35,6 Mill. Jahren datiert werden, da sie vor der Auffächerung der Kerngruppe der Familie Brassicaceae in die Abstammungslinien I und II stattgefunden hat. Die Auffächerung der Brassicaceae-Kerngruppe wird in nahe Verbindung mit einer Genomverdopplung (α) gebracht, welche vor ~ 40 Mill. Jahren (und früher) eintrat und neben dem

Wirtspflanzenwechsel der Familie der Pieridae (Beilstein et al. 2010) als Auslöser für die Auffächerung der Familie der Brassicaceae in Frage kommt (Franzke et al. 2011). Diese Tatsache würde zeitlich mit einer Diversifizierung der NSPs und der ESPs als Konsequenz der α -Genomduplikation übereinstimmen. Die α -Genomduplikation kann daher neben lokalen Genduplikationen gefolgt von einer Subneofunktionalisierung als möglicher Mechanismus dieser Diversifizierung in Betracht gezogen werden.

Nimmt man die α -Genomduplikation jedoch als Ursache der Verdopplung von At3g07720 und homologen Sequenzen oder der NSPs des AtNSP5-Clusters an, so kämen für die Entstehung der NSPs des AtNSP1-Clusters aus NSPs des AtNSP5-Clusters oder entsprechend der ESPs aus NSPs des AtNSP1-Clusters weitere nach der α -Genomduplikation datierte intragenomische Duplikationen von Chromosomen, Chromosomsegmenten und Genen in Frage, welche die Chance für die strukturellen Veränderungen der NSPs des AtNSP1-Clusters und für die Subneofunktionalisierung zu ESPs geliefert haben könnten.

Eine weitere bekannte Genomduplikation in der Ordnung Brassicales ist die ältere β -Genomduplikation. Sie trat bereits vor ~ 70 Mill. Jahren ein und könnte ein möglicher Auslöser für die Auffächerung der Kerngruppe der Ordnung Brassicales gewesen sein (Franzke et al. 2011). Da bisher keine spezifizierenden Proteine in Spezies außerhalb der Familie Brassicaceae gefunden werden konnten erscheint es unwahrscheinlich, dass bereits diese β -Genomduplikation zu einer Diversifizierung der spezifizierenden Proteine geführt haben könnte, kann aber nicht ausgeschlossen werden. Im Rahmen der vorliegenden Arbeit konnten zwar *ESP*-Sequenzen aus *C. spinosa* und *M. oleifera* isoliert werden (2.10), jedoch handelt es sich dabei vermutlich um Artefakte, so dass ein tatsächlicher Beweis für die Existenz von spezifizierenden Proteinen außerhalb der Familie der Brassicaceae aussteht. Sollte sich jedoch die Verdopplung von At3g07720 und homologer Sequenzen oder NSPs des AtNSP5-Clusters als Konsequenz der β -Genomduplikation beweisen lassen, wäre an dieser Stelle erneut die α -Genomduplikation neben lokalen Genduplikationen als Ursprung der Evolution der NSPs des AtNSP1-Clusters aus NSPs des AtNSP5-Clusters oder entsprechend der ESPs aus NSPs des AtNSP1-Clusters zu diskutieren.

Die abstammungsgeschichtlich jüngste Entwicklung innerhalb der spezifizierenden Proteine ist die Erlangung der Fähigkeit zur Formung organischer Thiocyanate der TFPs ausgehend von ESPs. TFPs konnten sowohl im Tribus Thlaspeideae als auch im Tribus Lepidieae identifiziert werden. Ihre zweifach unabhängige Entstehung konnte im Rahmen der vorliegenden Arbeit gezeigt werden. Auf Grund der Tatsache, dass sie unabhängig voneinander in den zwei Tribus und somit einmal in der Abstammungslinie I und einmal in der Abstammungslinie II entstanden sind, müssen sie nach der Auffächerung der Kerngruppe der Familie Brassicaceae in drei Abstammungslinien (I: vor ~ 35,6 Mill. Jahren; II: vor ~ 30,8 Mill. Jahren; III: vor ~ 35,5 Mill. Jahren) (Beilstein et al. 2010) aus den ESPs hervorgegangen sein. Als mögliche Mechanismen ihrer Entstehung kommen neben einer Subneofunktionalisierung ohne vorherige Genduplikation, also einer von Genduplikationen unabhängigen Entstehung, auch hier intragenomische Duplikationen von Chromosomen, Chromosomsegmenten und Genen in Frage. Eine weitere Eventualität stellen Tribus-spezifische Genomduplikationen dar. Diese Möglichkeit kommt auf Grund der

Tatsache in Betracht, dass TaTFP (Kuchernig et al. 2011) und ApTFP in zwei Spezies des Tribus Thlaspideae identifiziert werden konnten. Ihre Entwicklung muss somit vor Auffächerung dieses Tribus stattgefunden haben. Hinweise auf ähnliche Entwicklungen in der Evolution der Ordnung Brassicales stellen erst kürzlich aufgetretene Abstammungslinien-spezifische Genomduplikationen dar. So konnte für die Gattung *Brassica* eine Genomduplikation (Br- α) detektiert werden, welche der Auffächerung des Tribus Brassiceae voranging (Lagercrantz and Lydiat 1996; Franzke et al. 2011). Sie ist jünger als die α -Genomduplikation vor ~ 40 Mill. Jahren und fällt vermutlich in die Abstammungslinie II hinein. Desweiteren wurden zwei Genomduplikationen innerhalb der Abstammungslinie I identifiziert (Sn- α (Mandakova et al. 2010b; Franzke et al. 2011) und Pex- α (Mandakova et al. 2010a; Franzke et al. 2011)). Diese Duplikationen wurden auf 6 Mill. Jahre - 9 Mill. Jahre (Sn- α) sowie 0,8 Mill. Jahre bis 1,6 Mill. Jahre (Pex- α) datiert und haben voraussichtlich zusammen die Auffächerung und Mannigfaltigkeit des Tribus Camelinae beschleunigt (Franzke, Lysak et al. 2011). Eine sichere Aussage über mögliche Mechanismen der Evolution von TFPs ist jedoch nicht möglich, da TFPs im Rahmen der vorliegenden Arbeit nicht in Kombination mit ESPs oder NSPs innerhalb einer Spezies nachgewiesen werden konnten. Ihre Entstehung kann daher auch vollständig unabhängig von einer Genomduplikation erfolgt sein.

Die oben aufgeführten Schätzungen für ein evolutives Alter der spezifizierenden Proteine des Glucosinolat-Myrosinase-Systems wurden auf Basis von Veröffentlichungen über die Evolution innerhalb der Ordnung Brassicales und ihrer Entstehung vorgenommen. Hierbei ist zu beachten, dass allgemeine Altersangaben sehr kontrovers diskutiert werden. Bezogen wurde sich überwiegend auf Angaben in kürzlichen Veröffentlichungen von Beilstein et al. 2010 und Franzke et al. 2011.

3.4 Charakterisierung des TaTFP aus *T. arvense*

Die während der Charakterisierung des TaTFP (Kuchernig et al. 2011) erhaltenen Glucosinolatprofile von Blättern und Samen sowie das Auftreten von Allylglucosinolat als Hauptglucosinolat in *T. arvense* (Abbildung 21, Tabelle S 9) stimmen mit vorherigen Veröffentlichungen (Gmelin 1959; Lüthy and Benn 1977; Nielsen et al. 1979; Rodman and Chew 1980; Al-Shehbaz and Al-Shammary 1987) überein. Eine Gegenüberstellung der bereits im Vorfeld bekannten Merkmale des LsTFP (Burow et al. 2007a) zu den im Rahmen der vorliegenden Arbeit gewonnenen Ergebnissen für das TaTFP (2.11) zeigt deutlich das abweichende Glucosinolatprofil der untersuchten Spezies und das entsprechend veränderte Spektrum der entstehenden Glucosinolat-Hydrolyseprodukte (Tabelle 12). In *T. arvense* ist Allylglucosinolat das Hauptglucosinolat (Abbildung 21, Tabelle S 9) während das Hauptglucosinolat in *L. sativum* Benzylglucosinolat ist (Tabelle 12). Entsprechend der Hauptglucosinolate der jeweiligen Pflanzen formt das TaTFP Thiocyanat aus Allyl-, jedoch nicht aus Benzylglucosinolat (Abbildung 22). Hingegen ermöglicht LsTFP den Umsatz zu Thiocyanat aus Benzyl- jedoch nicht aus Allylglucosinolat (Tabelle 12). In Übereinstimmung mit den in der Pflanze vorhandenen Hauptglucosinolaten weisen TaTFP und LsTFP somit unterschiedliche Substratspezifitäten auf. Ein Einfluss der in den beiden Pflanzen auftretenden Glucosinolatprofile auf die Entwicklung der jeweiligen Substratspezifität der zweifach unabhängig entstandenen TFPs ist dabei sehr wahrscheinlich. Ohne das Vorhandensein eines umsetzbaren Substrates in einer Pflanze würde die Entwicklung eines spezifizierenden Proteins im Laufe der Evolution keinen selektiven Vorteil für die Pflanze bringen. Die eventuell in einer anderen Pflanze zu einer neuen Funktion führende Akkumulation an Mutationen würde sich hier als nutzlos erweisen und wieder „verloren gehen“. Dass TFPs jedoch nur in wenigen Pflanzen auftreten obwohl sowohl Benzyl- als auch Allylglucosinolat in vielen Pflanzen vorkommen zeigt deutlich das Zusammenspiel aller auf eine Pflanze wirkenden biotischen und abiotischen Faktoren und ihren gemeinsamen Einfluss auf die Evolution.

Das *TaTFP* wird in allen betrachteten ober- und unterirdischen Pflanzengeweben sowie generativen Organen (Wurzeln, Stängeln, Blättern, Blüten, unreifen Schoten sowie zwei Wochen alten Keimlingen) (Abbildung 24, Tabelle 12) transkribiert. Auch in *L. sativum* wird das *LsTFP* sowohl in ober- als auch in unterirdischen Organen (Wurzeln, Blättern, Blüten, Schoten sowie acht Tage alte Keimlinge) transkribiert. Hier zeigte die Expressionsanalyse die höchsten Transkriptlevel in Blüten (Burow et al. 2007a). Beiden Pflanzen gemein ist die Tatsache, dass die erhaltenen Transkriptionslevels in einzelnen Pflanzenorganen nicht den, auf Grund der erhaltenen Glucosinolat-Hydrolyseprodukte erwarteten, Transkriptionslevels entsprechen. Eine mögliche Erklärung wäre eine eventuelle posttranskriptionale Regulation der Expression. Für eine Aussage hierüber sind jedoch weitere Untersuchungen notwendig.

Weiterhin ist die Ähnlichkeit der Aminosäuresequenzen im Vergleich zwischen TaTFP und LsTFP untereinander deutlich geringer als jeweils im Vergleich zu den nächstähnlichen ESPs (Tabelle S 2, Abschnitt 2.6.1). Zusammengenommen mit der unterschiedlichen Substratspezifität der beiden Proteine stellt sich die Frage nach der strukturellen Basis der Substrat- und der Produktspezifität der beiden Proteine. Da sich TaTFP im Gegensatz zu

LsTFP in aktiver Form aufzureinigen lässt ist die Möglichkeit für Strukturanalysen durch Ermittlung der Kristallstruktur eine Option für zukünftige Untersuchungen des Proteins.

Tabelle 12: Gegenüberstellung der Eigenschaften von TaTFP, LsTFP und ihrer Herkunftsspezies. Aufgeführt sind die Glucosinolatprofile der Spezies *T. arvense* und *L. sativum* sowie die daraus hervorgehenden Glucosinolat-Hydrolyseprodukte nach Umsatz in einer Glucosinolat-Myrosinase-Reaktion in Anwesenheit von TaTFP oder LsTFP und das Expressionsprofil laut semiquantitativer PCR (Burow et al. 2007a; Kuchernig et al. 2011). GlS = Glucosinolat.

Merkmal	<i>T. arvense</i> TaTFP	<i>L. sativum</i> LsTFP
Glucosinolatprofil	> 90 % Allylglucosinolat Keimling, Stängel, Blatt, Blüte, Schote, Samen Max. 2 % Benzylglucosinolat Keimling	Hauptglucosinolat → Benzylglucosinolat Allylglucosinolat Samen
Spuren von:	Indol-3-ylmethylgls 1-Methoxy-indol-3-ylmethylgls 4-Methoxy-indol-3-ylmethylgls 4-Hydroxy-indol-3-ylmethylgls 3-Methylsulfinylpropylgls	4-Methoxy-indol-3-ylmethylgls 2-Phenylethylgls (Gil and MacLeod 1980)
Hydrolyseprodukte	Thiocyanat aus Allylgls Isothiocyanat aus Allylgls Nitril aus Allylgls Epithionitril aus Allylgls Nitril aus Benzylgls Isothiocyanat aus Benzylgls	Epithionitril aus Allylgls Isothiocyanat aus Allylgls Nitril aus Allylgls Thiocyanat aus Benzylgls Nitril aus Benzylgls Isothiocyanat aus Benzylgls
Expressionsprofil	In ober- und unterirdischen Pflanzenorganen	In ober- und unterirdischen Pflanzenorganen; in Blüten am höchsten

3.5 Ausblick

Wie bereits beschrieben gelang es im Rahmen der vorliegenden Arbeit den Datensatz der acht im Vorfeld bekannten spezifizierenden Proteine (Tabelle 1) auf 22 spezifizierende Proteine zu erweitern (Tabelle 6, Tabelle 7). Des Weiteren konnten sechs DOP-PCR-Fragmente mit Ähnlichkeit zu cDNAs spezifizierender Proteine erhalten werden (Tabelle 5, Tabelle 11). Eine funktionelle Charakterisierung dieser partiellen Sequenzen konnte nicht vorgenommen werden, jedoch wurde ihnen an Hand von Nukleotidsequenzähnlichkeiten zu bekannten spezifizierenden Proteinen mittels des EBI Webservers für ClustalW2 (Chenna et al. 2003) eine putative Funktion zugeordnet. Dies ergab die in Tabelle 13 angegebenen hypothetischen Eingruppierungen der Teilsequenzen.

Tabelle 13: Putative Funktionen der erhaltenen DOP-PCR-Fragmente. Angegeben sind Spezies und Zuordnungen der partiellen DOP-PCR-Fragmente. Die maximale Identität der Nukleotidsequenzen zu bekannten spezifizierenden Proteine entsprechend ClustalW (Chenna et al. 2003) in % ist gezeigt.

Spezies	Maximale Identität zu:	
<i>Alliaria petiolata</i>	76 %	<i>AtNSP1 / AtNSP2</i>
<i>Cardamine hirsuta</i>	74 %	<i>AtNSP2</i>
<i>Draba cinerea A</i>	92 %	<i>BrESP</i>
<i>Draba cinerea B</i>	73 %	<i>AtNSP2</i>
<i>Erysimum hieraciifolium</i>	76 %	<i>DaESP / TaTFP</i>
<i>Isatis tinctoria</i>	84 %	<i>AtNSP5</i>

Die Wahrscheinlichkeit der putativen Zuordnung der partiellen DOP-PCR-Fragmente konnte durch die Berechnung eines phylogenetischen Stammbaumes nach dem Neighbor Joining-Verfahren (Saitou and Nei 1987) erhöht werden (Abbildung 29). Die partiellen DOP-PCR-Fragmente ordnen sich hier entsprechend ihrer mittels ClustalW vorhergesagten maximalen Sequenzidentität ein (Abbildung 29). Hierbei bleibt die Auflösung des Stammbaumes nach Abstammungslinien erhalten. Zu beachten ist jedoch, dass eine Sequenzähnlichkeit keine sicheren Rückschlüsse auf die funktionelle Zuordnung eines Proteins zulässt. Diese ist nur durch biochemische Charakterisierung der *Full-length* Proteine zu erhalten. Aus diesem Grund wäre eine Vervollständigung der nur partiell erhaltenen DOP-PCR-Fragmente mittels 3'RACE-PCR und 5'RACE-PCR für weitere phylogenetische Analysen erstrebenswert.

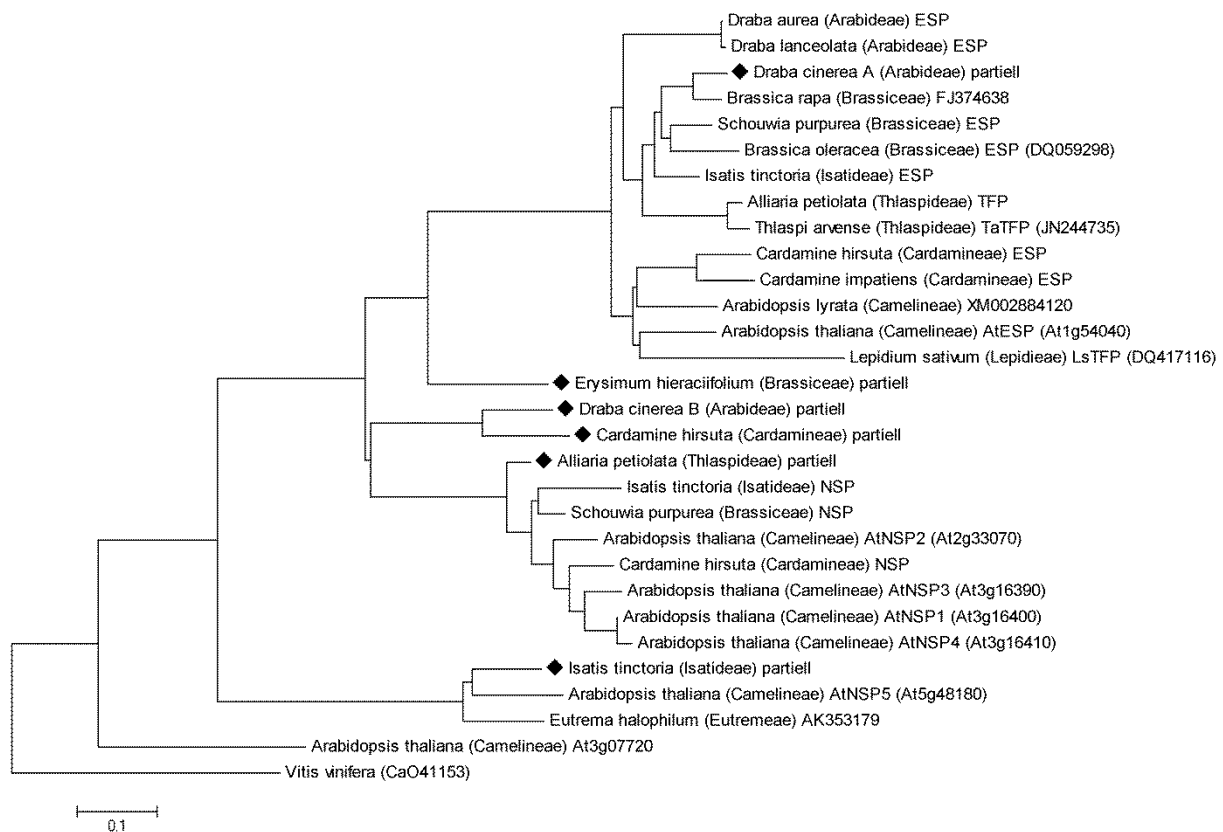


Abbildung 29: Neighbor Joining basierende Phylogenie spezifizierender Proteine und partieller DOP-PCR-Fragmente. Zur Erstellung verwendet wurden Nukleotidsequenzen spezifizierender Proteine der Ordnung Brassicales sowie die erhaltenen DOP-PCR-Fragmente nicht charakterisierter ähnlicher Proteine (◆). Der Neighbor Joining Baum basiert auf dem Kimura-2-Parameter-Modell und ist angegeben mit 1000facher Bootstrap Wiederholung. CAO41153 aus *V. vinifera* diente als Außengruppe zur Bewurzelung. Astlängen geben die geschätzten evolutiven Distanzen wieder. Die Maßstableiste zeigt die Länge für die angegebene Anzahl an Substitutionen pro Position.

Einen weiteren Ansatzpunkt für zukünftige Untersuchungen stellt die Identifizierung von Homologen von At3g07720 (Burow et al. 2009; Kissen and Bones 2009) und die Aufklärung ihrer biochemischen Funktion dar. Bis heute wurden At3g07720 und homologe Proteine aus der Ordnung Brassicales oder nicht-Glucosinolat-haltigen Pflanzen anderer Ordnungen nicht charakterisiert. Die bisher vorliegenden Anhaltspunkte für eine Funktion von At3g07720 in Wurzeln von *A. thaliana* in der Antwort auf Eisenmangel basieren ausschließlich auf einem erhöhten Transkriptionslevel des Proteins bei Eisenmangel (Colangelo and Guerinot 2004; Buckhout et al. 2009; Stein and Waters 2011). Die Tatsache, dass AtESP seine höchste ESP-Aktivität in Anwesenheit von 0,2 - 0,5 mM Eisenionen zeigt und auch eisenabhängig ist (Burow et al. 2006a), stellt einen interessanten Zusammenhang mit der möglichen Rolle von At3g07720 in der Antwort auf Eisenmangel dar. Andererseits scheint die Aktivität von TFPs und NSPs nicht von Eisenionen abzuhängen wird aber dadurch beeinflusst (Burow et al. 2009; Kuchernig et al. 2011). Eine Möglichkeit zur Charakterisierung von At3g07720 könnte die Analyse von Knockout-Mutanten bieten. Eine Beobachtung des Phänotyps dieser Pflanzen unter Eisenmangel bzw. Eisenfütterung könnte Aufschluss über eine tatsächliche Funktion von At3g07720 in der Antwort auf Eisenmangel liefern. Ein Vergleich des Eisengehaltes der Mutanten mit dem des Wildtyps könnte bei der Aufklärung der Beteiligung der Gene an der Eisenaufnahme helfen. Eine Fusion des Promotors von At3g07720 mit dem

Gen für eine β -Glucuronidase (GUS) oder dem „Green Fluorescent Protein“ (GFP) könnte nach dem Einbringen in das Genom von *A. thaliana* Aufschluss über seine Expression und deren Beeinflussungsmöglichkeiten bieten.

Im Rahmen der vorliegenden Arbeit wurden Hinweise auf das Vorhandensein von ESPs in den Spezies *C. spinosa* und *M. oleifera* gewonnen (2.10). Beide ESPs wurden auf Grund enormer Sequenzähnlichkeiten von 99 % untereinander und 98 % zu dem ESP aus *I. tinctoria* sowie der daraus resultierenden Vermutung einer eventuellen Verunreinigung nicht in die phylogenetischen Untersuchungen einbezogen. Beide Sequenzen liegen jedoch *Full-length* vor und wurden im Rahmen einer funktionellen Charakterisierung als ESPs bestätigt (2.10). Unter Berücksichtigung dieser Sequenzen würde sich die aufgestellte Entstehungsfolge der spezifizierenden Proteine auf einen erheblich früheren Zeitpunkt verschieben. Die Kerngruppe der Ordnung Brassicales entstand vor ~ 71,3 Mill. Jahren umfasst die Familien Brassicaceae, Capparaceae und Cleomaceae (Beilstein et al. 2010). Daher würde ein tatsächliches Vorhandensein spezifizierender Proteine in *C. spinosa* und *M. oleifera* die Entstehung der spezifizierenden Proteine auf > 71,3 Mill. Jahre datieren und somit in Verbindung mit der β -Genomverdopplung vor ~ 70 Mill. Jahren (Beilstein et al. 2010) bringen können. Im Southern-Kapillar-Blot konnte das Vorhandensein eines ESP in *C. spinosa* nicht aber in *M. oleifera* bestätigt werden (Abbildung 20B). Eine Bestätigung der Sequenzen und eine Identifizierung weiterer spezifizierender Proteine in Spezies außerhalb der Familie der Brassicaceae wären somit für weiterreichende Evolutionsanalysen innerhalb der spezifizierenden Proteine des Glucosinolat-Myrosinase-Systems wünschenswert.

Des Weiteren ist die Frage, ob es spezifizierende Proteine nur in Pflanzen der Brassicaceae gibt auch in Zusammenhang mit ihrer möglichen gemeinsamen Wirkung zusammen mit Nitrilasen im Katabolismus der Pflanzen sehr interessant. Zusätzlich zu ihrer Funktion in der Abwehrreaktion gibt es Hinweise darauf, dass sekundäre Metabolite wie die Glucosinolate auch einem endogenen Katabolismus unterworfen sind. Da Sekundärmetabolite eine Quelle für Stickstoff (z. B. cyanogene Glycoside) und / oder Schwefel (Glucosinolate) darstellen könnten, wäre ihr Recycling für die Pflanze zur Erhaltung ihrer Fitness sinnvoll (Janowitz et al. 2009). So wird zum Beispiel der hydrolytische Abbau des cyanogenen Glykosides Dhurrin zu Cyanid in *S. bicolor* von einem endogenen Katabolismus ergänzt, in welchem das Intermediat 4-Hydroxyphenylacetonitril durch einen Heterokomplex aus zwei Nitrilasen SbNIT4A / SbNITB2 zu 4-Hydroxyphenylelessigsäure und Ammoniak umgelenkt wird (Jenrich et al. 2007). Da Glucosinolate in einigen Pflanzengewebe 10 % - 30 % des pflanzlichen Schwefels enthalten (Falk et al. 2007), stellen sie eine mögliche Speicherform dieses Makronährstoffes dar (Janowitz et al. 2009). Für ein Recycling des Schwefels könnten die aus *A. thaliana* isolierten Nitrilasen NIT1 – NIT3 (X63445, U09958, U09959 (Bartling et al. 1992; Bartel and Fink 1994; Bartling et al. 1994)) dienen (Janowitz et al. 2009). Sie sind während der Evolution der Brassicaceae entstandene neofunktionalisierte Duplikate der Nitrilase NIT4 (U09961 (Bartel and Fink 1994)) und ihren Homologen und konnten auch in *B. rapa*, *B. napus* und *C. rubella* identifiziert werden (Janowitz et al. 2009). Die ursprüngliche Form der NIT4 und ihrer Homologen kommt in vielen Pflanzen vor und besitzt eine Funktion in der Detoxifikation von Cyanid. NIT1 – NIT3 haben eine Aminosäuresequenzidentität von 66 % - 68 % mit NIT4, sie haben jedoch die Funktion in der Detoxifikation von Cyanid

verloren. NIT1 und NIT2 aus *A. thaliana* sowie ihre Homologe in *B. rapa* und *C. rubella* besitzen jedoch eine breite Substratspezifität für aliphatische und aromatische Nitrile, welche der Glucosinolat-Myrosinase-Reaktion entstammen könnten. Diskutiert wird, ob die Evolution der Nitrilasen durch die Evolution der spezifizierenden Proteine angetrieben worden sein und zu einem ergänzenden katabolischen Glucosinolat-Abbauweg geführt haben könnte (Janowitz et al. 2009). In diesem Zusammenhang ist der dem Southern-Kapillar-Blot entstammende Hinweis auf das Vorhandensein einer *ESP*-ähnlichen Sequenz in *C. rubella* ein interessanter unterstützender Anhaltspunkt für diese Hypothese. Bisher ist der Nachweis von NIT1 Homologen nur in Spezies der Familie der Brassicaceae gelungen. Somit wäre eine erfolgreiche Klonierung von spezifizierenden Proteinen in Schwesterfamilien der Brassicaceae nicht nur eine interessante Ergänzung für die vorliegende phylogenetische Analyse der spezifizierenden Proteine, sondern könnte auch zu neuen Einsichten über Nitrilasen führen. Dies würde eine Kombination beider phylogenetischer Untersuchungen in der Gesamtbetrachtung der Evolution des Glucosinolat-Myrosinase-Systems ermöglichen.

4. ZUSAMMENFASSUNG

Das in der Ordnung Brassicales vorkommende Glucosinolat-Myrosinase-System ist eines der am besten untersuchten Systeme pflanzlicher Abwehr. Spezifizierende Proteine sind Teil dieses aktivierten chemischen Abwehrsystems. Bei Gewebeverletzung kommt es zur Hydrolyse der Glucosinolate durch Myrosinasen. Die entstehenden Glucosinolat-Hydrolyseprodukte lagern sich entweder spontan in toxische, als Abwehrstoffe dienende Isothiocyanate um oder sie werden durch spezifizierende Proteine in Epithionitrile, einfache Nitrile oder organische Thiocyanate mit weitgehend unbekannter biologischer Funktion umgewandelt. Ziel dieser Arbeit war die Klonierung, Charakterisierung und phylogenetische Analyse spezifizierender Proteine unterschiedlicher Produktspezifität.

Durch ein phytochemisches Screening innerhalb von sechs Familien der Ordnung Brassicales wurde der Nachweis für das Auftreten von, von Isothiocyanaten abweichenden, Glucosinolat-Hydrolyseprodukten in 14 Spezies der Familie der Brassicaceae sowie in einer Spezies der Familie Tropaeolaceae (*T. minus*) erbracht.

Ergänzend zu acht im Vorfeld bekannten spezifizierenden Proteinen aus *A. thaliana*, *B. oleracea* und *L. sativum* konnten *Full-length* cDNAs von drei weiteren NSPs, sechs ESPs und zwei TFPs isoliert und die codierten Proteine biochemisch charakterisiert werden. Insbesondere die Identifizierung des TFP aus *T. arvense* machte erstmals die Charakterisierung eines TFP in aufgereinigter Form möglich. Ferner wurden sechs DOP-PCR-Fragmente mit hoher Sequenzähnlichkeit zu spezifizierenden Proteinen erhalten, die jedoch nicht in die phylogenetischen Analysen einbezogen wurden. Für die phylogenetischen Analysen der spezifizierenden Proteine standen unter Einbeziehung von drei einer Datenbankrecherche entnommenen und nicht charakterisierten Proteinen 22 *Full-length* cDNAs und die homologe Sequenz At3g07720 aus *A. thaliana* zur Verfügung.

Die phylogenetischen Stammbäume wurden auf Basis der Nukleotid- und der Aminosäuresequenzen mittels des Neighbor Joining-, Maximum Likelihood- und des Maximum Parsimonie Verfahrens erstellt. Die grundlegend identische Topologie aller ermittelten phylogenetischen Stammbäume sowie die Berechnungen von Datenmatrizen für Sequenzdivergenzen und Sequenzidentitäten als Maß evolutionärer Distanz und die strukturellen Merkmale der spezifizierenden Proteine bestätigen At3g07720 als Vertreter einer Gruppe von Vorläuferproteinen der spezifizierenden Proteine. Ausgehend von At3g07720 und homologen Sequenzen entwickelten sich die NSPs des AtNSP5-Clusters. Eine Aufspaltung dieser ersten im Glucosinolat-Myrosinase-System funktionellen NSPs führte zu den NSPs des AtNSP1-Clusters. Eine darauf folgende Duplikation der NSPs des AtNSP1-Clusters führte dann zu den ESPs des ESP / TFP-Clusters. Die Auffächerung der ESPs konnte auf einen Zeitpunkt nach der Auffächerung der Kerngruppe der Familie Brassicaceae in die Abstammungslinien I und II geschätzt und die vorausgehende Genduplikation auf ein Alter von mindestens 35,6 Mill. Jahre datiert werden. In der weiteren evolutiven Entwicklung ergab sich eine im Rahmen der vorliegenden Arbeit erstmals identifizierte zweifach unabhängige Entstehung von TFPs ausgehend von ESPs, welche nach der Auffächerung der Kerngruppe der Familie Brassicaceae in die Abstammungslinien stattgefunden haben muss.

5. MATERIAL UND METHODEN

5.1 Material

5.1.1 Pflanzenmaterial

Die Herkunft der Samen ist Tabelle 14 zu entnehmen. Pflanzen wurden in Gewächshäusern des Arzneipflanzengartens des Instituts für Pharmazeutische Biologie oder in Kulturkammern (I-66LL) der Firma Percival Scientific, Inc. (Perry, USE) angezogen. Die verwendete Pflanzenerde Compo sana der Firma COMPO GmbH & Co. KG (Münster, Deutschland) wurde für 120 min bei 101°C gedämpft, bei 4°C gelagert und direkt vor Verwendung mit einer Verdünnung von 2 ml / 1 l Neudomück / H₂O (W. Neudorff GmbH KG, Emmerthal, Deutschland) durchtränkt. Samen wurden auf der behandelten Erdoberfläche ausgebracht und für drei Tage bei 4°C kaltgestellt. Anschließend wurden die Pflanzen bei einer Lichtintensität von 100 µmol m⁻² s⁻¹ und einer Luftfeuchtigkeit von 60 – 70 % kultiviert. Je nach gewählter Langtag- oder Kurztagwachstumsperiode waren die Pflanzen 10 std von 8.00 – 18.00 Uhr (Kurztag) oder 16 std von 7.00 – 23.00 Uhr (Langtag) dem Licht ausgesetzt. Geerntetes Pflanzenmaterial wurde entweder frisch verwendet oder bei - 80°C gelagert.

Tabelle 14: Herkunft der Samen zur Anzucht eingesetzter Pflanzen.

Spezies	Familie	Herkunft
<i>Alliaria petiolata</i>	Brassicaceae	Sanct Joseph, Missouri oder Waldwegrand Lahstedt / Groß Lafferde
<i>Alyssum alpestre</i>	Brassicaceae	Österr. Gartenbaugesellschaft Graz
<i>Arabidopsis thaliana Col-0</i>	Brassicaceae	Max-Planck-Institut für chemische Ökologie Jena
<i>Arabis glabra</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Barbarea intermedia</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Barbarea vulgaris</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Camelina microcarpa</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Capsella rubella</i>	Brassicaceae	Botanischer Garten Bochum
<i>Cardamine hirsuta</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Cardamine impatiens</i>	Brassicaceae	Botanischer Garten Tübingen
<i>Cleome spinosa</i>	Cleomaceae	Private Anzucht
<i>Diplotaxis muralis</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Draba aurea</i>	Brassicaceae	Botanischer Garten Bonn
<i>Draba cinerea</i>	Brassicaceae	Botanischer Garten Bonn

<i>Draba lanceolata</i>	Brassicaceae	Botanischer Garten Bonn
<i>Draba muralis</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Eruca sativa</i>	Brassicaceae	Arzneipflanzengarten TU Braunschweig
<i>Erysimum cheiri</i>	Brassicaceae	Botanischer Garten Caen
<i>Erysimum hieraciifolium</i>	Brassicaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Iberis amara</i>	Brassicaceae	Botanischer Garten Krefeld
<i>Isatis tinctoria</i>	Brassicaceae	Arzneipflanzengarten TU Braunschweig
<i>Limnanthes douglasii</i>	Limnanthaceae	Botanischer Garten Bonn
<i>Moringa oleifera</i>	Moringaceae	Botanischer Garten Nemmara-Palakkad
<i>Reseda lutea</i>	Resedaceae	Botanischer Garten Krefeld
<i>Reseda luteola</i>	Resedaceae	Botanischer Garten u. Rhododendrenpark Bremen
<i>Schouwia purpurea</i>	Brassicaceae	Arzneipflanzengarten TU Braunschweig
<i>Thlaspi arvense</i>	Brassicaceae	Arzneipflanzengarten TU Braunschweig
<i>Tropaeolum minus</i>	Tropaeolaceae	Botanischer Garten Gießen

5.1.2 Bakterienstämme und Plasmidvektoren

5.1.2.1 Plasmidvektoren

5.1.2.1.1 pGEM®-T-Easy (Promega)

Es handelt sich um einen 3015 bp großer linearisierter *high-copy* Vektor mit einem einzelnen 3'-terminalen Thymidin an beiden Enden. Das System dient der ungerichteten, einfachen Klonierung von PCR-Fragmenten welche mit einer Taq-Polymerase amplifiziert wurden (TA-Klonierung) und ermöglicht eine Blau-Weiß-Selektion.

5.1.2.1.2 pET-52b(+) (Novagen / Merck)

Dieser Expressionsvektor hat eine Größe von 5227 bp und seine Multiple-Cloning-Site ist N-terminal von einer Strep•Tag® II codierenden Sequenz (Skerra and Schmidt 2000) und einer Erkennungssequenz der 3C-Protease des humanen Retrovirus Typ 14 (HRV 3C Erkennungssequenz) (Cordingley et al. 1989) flankiert. C-terminal befindet sich eine His•Tag® codierende Sequenz (Hochuli et al. 1988). Das System ist durch Zusatz von Isopropyl-β-D-thiogalactopyranosid (IPTG) induzierbar.

5.1.2.1.3 pET-52b(+) ModI (AG Wittstock)

Der 5300 bp große Vektor stellt eine Modifikation des oben beschriebenen pET-52b(+) Vektors dar (AG Wittstock). Der originale pET-52b(+) Vektor wurde N-terminal der Multiple-Cloning-Region mit einer speziell generierten PacI-Kassette ausgestattet und ist mit der USER-Technik kompatibel (Nour-Eldin et al. 2006). Da die eingefügte PacI-Kassette N-terminal des vektoreigenen Strep-Tag des pET-52b(+) einkloniert wurde und dessen Leserahmen unterbricht, dient dieser Vektor der Expression und Funktionskontrolle von Proteinen ohne Möglichkeit zu deren Aufreinigung.

5.1.2.1.4 pET-52b(+) ModII (AG Wittstock)

Dieser Vektor hat eine Größe von 5378 bp großer Vektor und basiert auf der ersten Modifikation des Originalvektors pET-52b(+). Um auch bei Einsatz der USER-Technik die Möglichkeit zur Aufreinigung der exprimierten Proteine zu haben, wurde eine PacI-Kassette in pET-52b(+) ModI (AG Wittstock) eingebracht, die die Sequenz für einen N-terminalen Strep-Tag mit anschließender HRV 3C Erkennungssequenz beinhaltet (AG Wittstock). Unter Verwendung des Vektors pET-52b(+) ModII ist es möglich rekombinante Proteine mit N-terminal gelegenen Strep-Tag zu exprimieren.

5.1.2.2 Bakterienstämme

5.1.2.2.1 *E. coli* XL1Blue (Agilent Technologies)

Dieser Bakterienstamm mit dem Genotyp *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacIq* ZΔM15 Tn10 (Tetr)] wurde für anfängliche Routineklonierungen sowie das Anlegen von Glycerinkulturen genutzt.

5.1.2.2.2 *E. coli* BL21(DE3) pLysS (Invitrogen)

Dieser Bakterienstamm mit dem Genotyp F- *ompT hsdSB(rB-,mB-) gal dcm* (DE3) pLysS (CamR) wurde für die Expression rekombinanter Proteine genutzt.

5.1.3 Primer

Die in Tabelle S 10 aufgelisteten Primer wurden für die Identifizierung der in Tabelle 4 aufgelisteten cDNAs verwendet. Sie weisen eine Länge von 18 – 25 Nukleotiden und einen GC-Gehalt von ~ 50 % auf. Ihre Schmelztemperatur liegt 50°C - 65°C. Alle Primer wurden von der Firma Invitrogen (Invitrogen by Life Technologies Corporation) bezogen, zu 100 pmol / µl in 10 mM Tris-HCl-Puffer (pH 8,0) gelöst und bei -20°C gelagert. Degenerierte Primer enthielten entweder Mischbasen oder Inosin. Inosin vermag mit Adenin (A), Guanin (G), Cytosin (C) und Thymin (T) zu hybridisieren (Case-Green and Southern 1994) und wurde eingesetzt um einen zu hohen Degenerationsgrad zu vermeiden. Der Degenerationsgrad berechnet sich als Produkt der Anzahl möglicher Basen in jeder degenerierten Position.

5.1.4 Enzyme

– SuperScript® III RT	Invitrogen by Life Technologies Corporation Carlsbad, Kalifornien
– BioScript™	Bioline GmbH Luckenwalde, Deutschland
– peqGOLD <i>Taq</i> DNA Polymerase	PEQLAB Biotechnologie GMBH Erlangen, Deutschland
– <i>Pfu</i> DNA Polymerase (recombinant)	Fermentas GmbH St. Leon-Rot, Deutschland
– omni Taq -OLS®-DNA-Polymerase	OLS OMNI Life Science GmbH & Co. KG Bremen, Deutschland
– Mol Taq DNA-Polymerase	Molzym GmbH & Co. KG Bremen, Deutschland
– <i>Pfu</i> Turbo Cx Hotstart DNA Polymerase	Agilent Technologies Waldbronn, Deutschland
– Terminal Deoxynucleotidyl Transferase (TdT)	Fermentas GmbH St. Leon-Rot, Deutschland

- | | |
|---------------------------------------|--|
| – RNase H, <i>E. coli</i> | Fermentas GmbH
St. Leon-Rot, Deutschland |
| – <i>PacI</i> (10.000 Units / ml) | New England Biolabs GmbH
Frankfurt am Main, Deutschland |
| – <i>Nt.BbvCI</i> (10.000 Units / ml) | New England Biolabs GmbH
Frankfurt am Main, Deutschland |
| – <i>EcoRI</i> (10 Units / µl) | Fermentas GmbH
St. Leon-Rot, Deutschland |
| – <i>NcoI</i> (19 Units / µl) | Fermentas GmbH
St. Leon-Rot, Deutschland |
| – <i>EcoRI</i> , HC (50 Units / µl) | Fermentas GmbH
St. Leon-Rot, Deutschland |
| – <i>BamHI</i> , HC (50 Units / µl) | Fermentas GmbH
St. Leon-Rot, Deutschland |

5.1.5 Kits

- | | |
|--|---|
| – ribozo/OLS® | OLS OMNI Life Science GmbH & Co. KG
Bremen, Deutschland |
| – innuPREP Plant RNA Kit | Analytikjena
Jena, Deutschland |
| – QIAEX II Gel Extraction Kit | Germany — QIAGEN GmbH
Hilden, Deutschland |
| – Wizard® Plus SV Minipreps
DNA Purification System | Promega GmbH
Mannheim, Deutschland |
| – Strep-tag® purification system | IBA GmbH
Göttingen, Deutschland |
| – Pierce® BCA Protein Assay Kit | Thermo Fisher Scientific
Bonn, Deutschland |
| – PCR DIG Probe Synthesis Kit | Roche Diagnostics Deutschland GmbH
Mannheim, Deutschland |
| – DIG-High Prime DNA Labeling
Detection Starter Kit I | Roche Diagnostics Deutschland GmbH and
Mannheim, Deutschland |

5.1.6 Weitere Reagenzien und Materialien

- | | |
|---|--|
| – dNTP Mix, 10 mM each | Fermentas GmbH
St. Leon-Rot, Deutschland |
| – dATP, 100 mM | Fermentas GmbH
St. Leon-Rot, Deutschland |
| – 1 kb DNA Ladder | Invitrogen by Life Technologies Corporation
Carlsbad, Kalifornien |
| – 100 bp DNA Ladder | Invitrogen by Life Technologies Corporation
Carlsbad, Kalifornien |
| – Fast-Ruler DNA Ladder | Fermentas GmbH
St. Leon-Rot, Deutschland |
| – Nylonmembran (10 x 15 cm),
positiv geladen | Roche Diagnostics Deutschland GmbH
Mannheim, Deutschland |
| – Whatman-Filterpapiere | CARL ROTH GmbH + Co. KG
Karlsruhe, Deutschland |

5.2 Methoden

5.2.1 Molekularbiologische Methoden

5.2.1.1 RNA-Isolierung

Die Isolation von RNA wird z.B. genutzt um Template-RNA für die nachfolgende cDNA-Synthese zu erhalten. Eine mögliche Methode ist die Verwendung von ribozol-OLS® (OLS) und beruht auf einem Protokoll von Chomczynski und Sacchi (Chomczynski and Sacchi 1987). Eine weitere Methode bietet das innuPREP Plant RNA Kit (Analytikjena). Die Isolation erfolgte laut der jeweiligen Herstellerangaben.

5.2.1.2 cDNA-Synthese

cDNA wird durch das Enzym Reverse Transcriptase (RT) aus RNA gewonnen und dient z.B. als Template für Genexpressionsanalysen und für die Gewinnung von DNA-Sequenzen ohne Intron-Exon-Strukturen. Im Rahmen dieser Arbeit wurde die RT SuperScript® III Reverse Transcriptase (Invitrogen) verwendet. Sie stellt eine Ribonuclease (RNase) H-reduzierte und hitzestabile Mutante der Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV) (Roth et al. 1985) dar. Eine zweite verwendete Variante der Reversen Transkriptase, ebenfalls eine RNase H-reduzierte Mutante der M-MLV, ist die Reverse Transkriptase BioScript™ (Bioline). Diese Variante wurde ausschließlich zur Expressionsanalyse in *T. arvense* angewandt.

5.2.1.2.1 cDNA-Synthese mit SuperScript® III RT (Invitrogen)

- 1 µl Oligo (dT)₁₂₋₁₈ (10 pmol / µl) und 1 µg RNA wurden mit ddH₂O auf 13 µl aufgefüllt.
- Nach Zugaben von 1 µl dNTP Mix (10 mM jedes) wurde für 5 min bei 65°C erhitzt.
- Anschließend wurde für 1min auf Eis inkubiert und anschließend kurz abzentrifugiert um eventuelles Kondenswasser zu sammeln.
- Es wurden 4 µl 5 x First-Strand-Buffer, 1 µl 0,1 M Dithiothreitol (DTT) und 1 µl SuperScript® III RT zugegeben und vorsichtig gemischt.
- Es folgten zwei Inkubationsschritte für 45 min bei 50°C und für 15 min bei 70°C.

5.2.1.2.2 cDNA-Synthese mit BioScript™ (Bioline)

- 0,1 µg, 0,3 µg und 1 µg RNA sowie 0,5 µg Oligo (dT)₁₈ wurden mit ddH₂O auf 12 µl aufgefüllt.
- Nach einer Inkubation von 5 min bei 70°C wurde die Probe für 2-3 min auf Eis gestellt.
- Anschließend wurde 1 µl dNTP Mix (10 mM jedes) sowie 4 µl 5 x Puffer zugegeben und auf ein Gesamtvolumen von 19,75 µl mit ddH₂O aufgefüllt.
- Nach vorsichtigem Mischen wurden 0,5 µl BioScript™ zugesetzt und für 60 min bei 41°C inkubiert.
- Ein zweiter Inkubationsschritt erfolgte bei 70°C für 10 min mit anschließender Lagerung auf Eis.

5.2.1.3 PCR-Techniken

5.2.1.3.1 Standard-PCR

Die PCR (Mullis et al. 1986) ist eine molekularbiologische Methode die der *in vitro* Herstellung großer Mengen definierter DNA-Fragmente dient. Als Polymerase wird sehr häufig die Taq-Polymerase, eine thermostabile DNA-Polymerase der Bakterienart *Thermophilus aquaticus* eingesetzt. Diese Polymerase bietet eine hohe Amplifikationsgeschwindigkeit und hängt durch ihre Template-unabhängige terminale Transferaseaktivität unspezifisch einzelne Adenosinphosphate an die 3'-Termini des PCR-Produktes an. Diese Besonderheit erlaubt die Ligation mit einem linearisierten Vektor der einen einfachen Thymidinphosphat-Überhang besitzt (z.B. pGEM®-T-Easy, 5.1.2.1.1). Taq-Polymerasen weisen jedoch keine *proof reading*-Funktion auf, daher ist zu beachten, dass sie fehlerhaft arbeiten können. Weniger fehlerhaft arbeiten Pfu-Polymerase aus der Archaeobakterienart *Pyrococcus furiosus*. Sie besitzt 3'→5' Exonukleaseaktivität und sind somit in der Lage, eingetragene Nukleotidfehler zu korrigieren. Eine terminale Transferaseaktivität besitzen sie nicht. Ein Adenosinphosphat-Überhang muss für eine TA-Klonierung nachträglich erzeugt werden.

Typischerweise wird das Template mit zwei das zu amplifizierende Fragment flankierenden Primern, Polymerasepuffer, dNTP Mix, ddH₂O und Polymerase inkubiert. Nach einem anfänglichen initiierenden Denaturierungsschritt bei 94°C – 95°C (3 min – 5 min) folgt jeder von ca. 30 – 40 PCR-Zyklen dem nachstehenden Verlauf:

- Denaturierung bei 94°C – 95°C (30 sek – 60 sek)
- Primer-Annealing bei 55°C – 60°C (30 sek – 60 sek)
- Elongation bei 70°C – 75°C (30 sek – 60 sek)

Ein abschließender Syntheseschritt bei entsprechender Elongationstemperatur für 5 min - 10 min stellt sicher, dass alle amplifizierten Fragmente komplementiert werden können.

Im Rahmen dieser Arbeit wurden, je nach Verfügbarkeit, Thermocycler der Firmen Biometra GmbH, Göttingen, Deutschland (TProfessional Basic und TProfessional Basic-Gradient), Eppendorf AG, Wesseling / Berzdorf, Deutschland (Mastercycler personal) und PEQLAB Biotechnologie GMBH, Erlangen, Deutschland (peqSTAR 96 Universal) verwendet.

5.2.1.3.2 Semiquantitative PCR zur Expressionsanalyse

Die semiquantitative PCR entspricht einer Standard-PCR (5.2.1.3.1) und wurde für die Expressionsanalyse in *T. arvense* eingesetzt. Als Template diente je 1 µl der durch die Reverse Transcriptase BioScript™ (Bioline) aus 0,1 µg, 0,3 µg und 1 µg RNA gewonnenen cDNAs (5.2.1.2.2). Je cDNA wurde sowohl ein PCR-Ansatz unter Verwendung von genspezifischen Primern als auch unter Verwendung von *Actin8*-spezifischen Primern durchgeführt. Für diese PCR-Technik wurde die peqGOLD Taq DNA Polymerase (PEQLAB) nach folgendem Schema eingesetzt:

Zusammensetzung des PCR-Ansatzes:

+	1	µl Template
+	1	µl dNTP Mix (10 mM jedes)
+	5	µl 10 x Puffer
+	0,25	µl peqGOLD Taq DNA Polymerase
+	1	µl Primer A (10 pmol / µl)
+	1	µl Primer B (10 pmol / µl)
+	40,75	µl ddH ₂ O
Σ 50		µl

PCR-Programm:

94 °C	3 min	} 35 x
94 °C	75 sek	
61,5°C	150 sek	
72 °C	1 min	
72 °C	7 min	
15 °C	∞	

5.2.1.3.3 RACE-PCR

Die RACE-PCR ist eine PCR-Technik, mit deren Hilfe unbekannte 5'- und 3'-Sequenzen ausgehend von einem bekannten Fragment, identifiziert werden können (Frohman et al. 1988).

Für eine 3'RACE-PCR bedient man sich dem am 3'Ende der mRNA durch das Enzym Poly(A)-Polymerase angehängten Poly(A)-Schwanz, um die mRNA mit Hilfe eines Oligo (dT)-Anker-Primers sowie der RT in cDNA zu transkribieren. Das so entstandene Template wird dann für eine PCR-Reaktion mit einem genspezifischen Primer (GSP1) genutzt, welcher an dem bekannten Fragment hybridisiert und so orientiert ist, dass das gesuchte Ende zwischen dem genspezifischen und dem erneut eingesetzten Oligo (dT)-Anker-Primer amplifiziert werden kann. Eine nested-PCR mit einem weiteren genspezifischen Primer (GSP2) und einem Anker-Primer ist möglich. Eine nested PCR beschreibt eine PCR-Technik, bei welcher zwei PCR-Ansätze nacheinander vorgesehen sind, um die Spezifität der Amplifikation zu erhöhen. Dies geschieht durch den Einsatz eines Aliquot der ersten PCR-Reaktion als Template für die darauf folgende zweite PCR-Reaktion. In diesem zweiten PCR-Ansatz wird ein weiteres Primerpaar eingesetzt, welches innerhalb der, in der ersten PCR-Reaktion amplifizierten, Matrize bindet. Ein kürzeres PCR-Fragment mit erhöhter Spezifität wird amplifiziert.

cDNA Erststrangsynthese mit SuperScript® III RT (Invitrogen):

- 1 µl Oligo (dT)₁₈-Anker, 1 µl dNTP Mix (10 mM jedes) und 1 µg RNA wurden mit ddH₂O auf eine Gesamtvolumen von 13 µl aufgefüllt.
- Inkubation für 5 min bei 65°C gefolgt von einer Inkubation für 1 min auf Eis.
- Zugabe von 4 µl 5 x First-Strand-Buffer, 1 µl 0,1 M DTT und 1 µl SuperScript® III RT
- Inkubation für 45 min bei 50°C und 15 min bei 70°C
- Nach Zugabe von 1 µl RNase H wird für 20 min bei 37°C inkubiert.

Für die folgende PCR-Reaktion wurden je nach Template eingesetzt:
 peqGOLD Taq DNA Polymerase (A) oder Pfu DNA Polymerase (recombinant) (B).

Zusammensetzung des PCR-Ansatzes (A):

–	1	µl Template
+	0,5	µl dNTP Mix (10 mM jedes)
+	5	µl 10 x Puffer
+	4	µl MgCl ₂ (25 mM)
+	0,5	µl DNA-Polymerase
+	1	µl Primer A (10 pmol / µl)
+	1	µl Primer B (10 pmol / µl)
+	37	µl ddH ₂ O
<hr/>		
Σ	50	µl
<hr/>		

Zusammensetzung des PCR-Ansatzes (B):

–	1	µl Template
–	0,5	µl dNTP Mix (10 mM jedes)
–	10	µl 5 x First-Strand-Buffer
–	4	µl MgCl₂ (25 mM)
–	0,5	µl DNA-Polymerase
–	1	µl Primer A (10 pmol / µl)
–	1	µl Primer B (10 pmol / µl)
–	36	µl ddH ₂ O
<hr/>		
Σ	50	µl
<hr/>		

PCR-Programm:

94°C	3	min	} 35 x
94°C	75	sek	
Xx°C	150	sek	
72°C	1	min	
<hr/>			
72°C	10	min	
15°C	∞		

Eine nested-PCR erfolgte nach jeweils gleichem Pipettierschema und mit gleichem PCR-Programm unter Einsatz von 1 µl des vorangegangenen PCR-Ansatzes als Template.

Für eine 5'RACE-PCR wird die mRNA unter Zuhilfenahme eines antisense gerichteten genspezifischen Primers (GSP1) mit RT in cDNA transkribiert. So erzielt man das Umschreiben nur spezifischer mRNA in cDNA, welche als Template für das Enzym Terminale Deoxynucleotidyltransferase (TdT) dient. Dieses Enzym ist temperaturunabhängig und katalysiert die Addition von Deoxynucleotiden an das 3'Ende von DNA-Molekülen. Eine so unter Einbau von Adenosinphosphat geteilte cDNA wird dann für eine PCR-Reaktion mit einem zweiten antisense Primer (GSP2) in Kombination mit einem Oligo (dT)-Anker-Primer genutzt. So kann das unbekannte Fragment zwischen genspezifischem Primer und 5'Ende der Sequenz amplifiziert werden. Eine zweite nested-PCR mit einem weiteren genspezifischen Primer (GSP3) und einem Anker-Primer ist möglich.

cDNA Erststrangsynthese mit SuperScript® III RT (Invitrogen):

- 2 µl GSP1, 1 µl dNTP Mix (10 mM jedes) und 1 µg RNA wurden mit ddH₂O auf 15 µl aufgefüllt.
- 5 µl 5 x First-Strand-Buffer und 3 µl 0,1 M DTT wurden auf 50°C vorgewärmt und nach einer Inkubation für 10 min bei 70°C zugegeben.
- Der Ansatz wurde für 2 min bei 50°C inkubiert.
- 1 µl SuperScript® III RT wurde hinzugefügt gefolgt von zwei Inkubationsschritten für 60 min bei 55°C und für 15 min bei 70°C.
- Nach Zugabe von 1 µl RNase H Inkubation für 45 min bei 50°C und 15 min bei 70°C
- Nach Zugabe von 1 µl RNase H wurde für 20 min bei 37°C inkubiert.

TdT-Tailing mit Terminaler Deoxynucleotidyltransferase (Fermentas):

- 10 µl Template wurden mit 5 µl 5 x Tailing Puffer, 2,5 µl dATP (100 mM) und 6,5 µl ddH₂O für 2-3 min bei 94°C inkubiert
- Nach einer Inkubation für 1 min auf Eis wurde 1 µl TdT zugegeben und 10 min bei 37°C gefolgt von 10 min bei 65°C inkubiert.

Die anschließende PCR-Reaktion erfolgte nach jeweils gleichen Bedingungen, wie sie für die 5'RACE-PCR beschrieben wurden.

5.2.1.3.4 Full-length-PCR

Diese Form der PCR-Reaktion entspricht einer Standard-PCR (5.2.1.3.1) unter Verwendung von zwei möglichst weit 5'- und 3'-terminal gelegenen genspezifischen Primern und unter ausschließlichem Einsatz einer Polymerase mit 3'→5' Exonukleaseaktivität, die Fehler bei der DNA-Synthese minimiert (hier *Pfu* DNA Polymerase (Fermentas)). Die *Full-length*-PCR wird verwendet, um den vollständigen ORF mit oder ohne untranslatierte Regionen (UTR) für Sequenzbestätigungen zu amplifizieren.

Zusammensetzung des PCR-Ansatzes:

+	1	µl Template
+	1	µl dNTP Mix (10 mM jedes)
+	5	µl 10 x <i>Pfu</i> Puffer mit MgSO ₄
+	0,5	µl <i>Pfu</i> DNA Polymerase
+	1	µl je Primer (10 pmol / µl)
+	40,5	µl ddH ₂ O
<hr/>		
Σ	50	µl
<hr/>		
<hr/>		

PCR-Programm:

95 °C	2	min	} 35 x
95 °C	1	min	
Xx °C	1	min	
72,5°C	1	min	
72,5°C	10	min	
15 °C	∞		

5.2.1.3.5 DOP-PCR

Diese PCR-Technik unterscheidet sich nur durch den Einsatz degenerierter Primer von einer Standard-PCR. Für die DOP-PCR wurde die *peqGOLD Taq*-DNA-Polymerase (PEQLAB) eingesetzt, eine thermostabile 5'→3' DNA-Polymerase für Standard-PCR-Amplifikationen, die eine nur schwache 5'→3'-Exonukleaseaktivität und keinerlei 3'→5'-Exonukleaseaktivität aufweist.

Zusammensetzung des PCR-Ansatzes:

+ 1,6 µl Template
+ 0,5 µl dNTP Mix (10 mM jedes)
+ 5 µl 10 x Puffer
+ 4 µl MgCl ₂ (25 mM)
+ 0,5 µl <i>peqGOLD Taq</i> -DNA-Polymerase
+ 1 µl Primer A (10 pmol / µl)
+ 1 µl Primer B (10 pmol / µl)
+ 37 µl ddH ₂ O
Σ 50 µl

PCR-Programm:

94°C	3 min	} 35 x
94°C	75 sek	
54°C – 58°C	150 sek	
72°C	1 min	
72°C	10 min	
15°C	∞	

5.2.1.3.6 Colony-screening-PCR

Eine Colony-screening-PCR ist eine Variation der Standard-PCR (5.2.1.3.1), bei welcher DNA transformierter *E. coli*-Zellen als Template dient, ohne dass diese erst isoliert werden muss. In der Colony-Screening-PCR werden Primer verwendet, welche gesuchte DNA-Fragmente flankieren. Zum Einsatz kommt die Colony Screening-PCR als schnelle und kostengünstige Methode zur Durchsicht und Auswahl von Klonen.

5.2.1.4 Agarose-Gelelektrophorese

DNA-Moleküle lassen sich aufgrund ihrer negativen Eigenladung ihrer Größe nach durch Gelelektrophorese auftrennen. Die Detektion der DNA im ultravioletten (UV) -Licht (366 nm) erfolgte durch Zugabe von Ethidiumbromid (0,25 µg / ml). Der Farbstoff wurde nach Aufkochen der Agarose in 1 x Tris-Borat-EDTA-Puffer (TBE) zugegeben und die Flüssigkeit auf Gelträgern mit Taschenkämmen ausgekühlt. DNA-Proben wurden mit einem 6 x Ladepuffer versetzt und aufgetragen. Die Auftrennung erfolgte in einem 0,7 - 2 %igen Agarose-Gel in 1 x Tris-Borat-EDTA-Puffer (TBE-Puffer) bei einer Spannung von 5 - 10 Volt pro cm Gellänge. Verwendete Gelkammern, -träger und -kämme wurden von der Firma Bio-Rad Laboratories GmbH (München, Deutschland) bezogen. An Hand des mitgeführten Standards bekannter Größen (1 kb DNA Ladder, 100 bp DNA Ladder, Invitrogen) konnte eine Größenzuordnung erfolgen.

5.2.1.5 Isolation von DNA-Fragmenten aus einem Agarose-Gel

Nach Auftrennung der DNA-Fragmente in einem Agarose-Gel und Detektion im UV-Licht wurden gewünschte Banden aus dem Agarose-Gel ausgeschnitten. Für die weitere Verwendung der enthaltenen DNA-Fragmente ist es notwendig, diese aus der Gelmatrix zu eluieren. Die DNA-Isolierung aus dem Agarose-Gel wurde mit Hilfe des QIAEX II Gel Extraction Kit (Qiagen) vorgenommen und folgte der Anleitung des Herstellers. Das Prinzip der Elution basiert auf dem Lösen der Agarose und der selektiven Bindung der DNA-Fragmente an die QIAEX II Silica-Gel Partikel in Anwesenheit chaotroper Salze.

5.2.1.6 Ligation von DNA-Fragmenten

Die kovalenten Phosphodiester-Bindungen eines DNA-Stranges können durch Restriktionsenzyme an spezifischen Stellen gespalten werden. Es entstehen DNA-Fragmente unterschiedlicher, definierter Größe. Auf diese Weise erzeugte DNA-Fragmente mit kompatiblen Enden (kohäsiv bzw. blunt) können neukombiniert (ligiert) werden. Durch Knüpfung der 3', 5'-Phosphodiesterbindung unter ATP-Hydrolyse, eine Reaktion der T4-DNA-Ligase (Promega), werden dabei die linearen DNA-Fragmente kovalent miteinander verbunden (Engler 1982). Eine einfachere und im Rahmen dieser Arbeit überwiegend verwendete Variante ist die TA-Klonierung mit dem pGEM®-T-Easy Vektor-System (5.1.2.1.1) nach Herstellerangaben.

5.2.1.7 Transformation kompetenter *E. coli*-Zellen

Für die Amplifikation rekombinanter Plasmid-DNA wurde diese in kompetente *E. coli*-Zellen des Stammes XL1Blue (Agilent Technologies) (5.1.2.2.1) eingebracht (transformiert). Die im Rahmen dieser Arbeit verwendete Methode zur Transformation beruhte auf der Erzeugung von Membranporen durch einen kurzen Hitzeschock und der Fähigkeit von kompetenten Bakterienzellen zur Aufnahme von Kationen-DNA-Komplexen. Hierfür wurde der vollständige Ligationsansatz (5.2.1.12) oder bei Verwendung aufgereinigter Plasmid-DNA (5.2.1.11) ein 100 ng Plasmid-DNA enthaltendes Aliquot verwendet. 30 - 50 µl kompetente *E. coli*-Zellen wurden auf Eis aufgetaut und mit der Plasmid-DNA vorsichtig auf Eis gemischt. Nach 30 min Inkubation auf Eis wurde das Gemisch für 45 sek einem Hitzeschock bei 42°C im Wasserbad ausgesetzt. Anschließend folgte eine sofortige Abkühlung für 2 min auf Eis und eine Zugabe von 250 µl Super Optimal Broth-Medium (SOC-Medium). Nach einer Inkubation von 45 min bei 37°C wurde der Ansatz auf LB-Platten (Lysogeny-Broth-Medium auch Luria-Bertani-Medium, (Bertani 1951)), versetzt mit 25 µl Ampicillin (Amp) (100 µg / ml), 70 µl (5-Bromo-4-Chloro-3-Indolyl-Beta-D-Galactopyranoside) (X-Gal) (2 %ig) und 65 µl IPTG (1 M) pro 25 ml Medium, ausgestrichen und über Nacht bei 37°C inkubiert. Für eine spätere Expression rekombinanter Proteine wurde der *E. coli*-Stamm BL21(DE3) pLysS (Invitrogen) (5.1.2.2.2) gewählt. Die Selektionsfaktoren wurden je nach Vektor angepasst.

5.2.1.8 Herstellung kompetenter *E. coli*-Zellen

Für die Transformation kompetenter *E. coli*-Zellen müssen diese Zellen für die verwendete Transformationsmethode vorbehandelt werden. Im Fall der Hitzeschock-Transformation wurde für die *E. coli* Stämme XL1Blue (Agilent Technologies) (5.1.2.2.1) und BL21(DE3) pLysS (Invitrogen) (5.1.2.2.2) die Calciumchloridmethode (Mandel and Higa 1970) verwendet.

- Aus einer bei -80°C gelagerten und auf Eis aufgetauten Kyrkultur wurde ein Aliquot auf einer LB-Platte ausgestrichen.
- Nach Inkubation über Nacht bei 37°C wurde eine 5 ml LB-Flüssigkultur mit einer Einzelkolonie angeimpft und ebenfalls über Nacht bei 37°C schüttelnd inkubiert.
- 50 ml LB-Flüssigmedium (250 ml Erlenmeyerkolben) wurden mit 500 µl der Über-Nacht-Flüssigkultur angeimpft und schüttelnd bei 37°C bis zu einer optischen Dichte (OD) OD₆₀₀ von 0,4 - 0,5 angezogen.
- Die Kultur wurde in ein 50 ml Falconröhrchen überführt und nach 10 min Abkühlen auf Eis für 10 min bei 4°C und 3000 rpm zentrifugiert.
- Das Pellet wurde anschließend in 25 ml Calciumchloridlösung (0,1 M) resuspendiert und 10 min auf Eis inkubiert.
- Nach einer Zentrifugation von 5 min bei 4°C und 3000 rpm wurde das Pellet in 5 ml eiskalter, mit 15 % Glycerin versetzter Calciumchloridlösung (0,1 M) aufgenommen und in Aliquotes bei -80°C gelagert.

5.2.1.9 Kyrkulturen von *E. coli*-Zellen

Wichtige Bakterien-Klone, die für spätere Arbeiten zur Verfügung stehen sollten, wurden kryokonserviert. Dazu wurden 750 µl einer LB-Flüssigkultur mit 250 µl LB-Glycerin-Medium (60 %ig) versetzt und gut vermischt. Die entstandene Bakterien-Suspension wurde nach dem Mischen sofort in flüssigem Stickstoff eingefroren und bei -80°C gelagert.

5.2.1.10 Uracil-Specific Excision Reagent (USER)-Klonierung

5.2.1.10.1 Vorbereitung der Vektoren

Da die Vektoren pET-52b(+) ModI und ModII (AG Wittstock) nicht linearisiert vorlagen mussten sie durch einen jeweiligen Restriktionsverdau für die USER-Klonierung vorbereitet werden. Nach einer Inkubation des Restriktionsansatz1 über Nacht bei 37°C wurden diesem am nächsten Morgen die unter Restriktionsansatz2 angegebenen Komponenten beigemischt:

Restriktionsansatz1:

+	10	µl pET-52b(+) ModI Vektor (1,4 µg)
+	5	µl 10 x NEB 4 Puffer
+	5	µl 10 x BSA
+	1,5	µl <i>PacI</i> (15 Units)
+	28,5	µl ddH ₂ O
Σ 50		µl

Restriktionsansatz2:

–	50	µl Restriktionsansatz1
–	0,5	µl 10 x NEB 4 Puffer
–	0,5	µl 10 x BSA
–	1	µl <i>PacI</i> (10 Units)
–	0,75	µl <i>Nt.BbvCI</i> (7,5 Units)
	2,25	µl ddH ₂ O
Σ 55		µl

Nach einer Inkubation für 2 std bei 37°C wurde der Ansatz in einem 1 %igen Agarose-Gel aufgetrennt (5.2.1.4) und entsprechende Banden ausgeschnitten. Die DNA wurde mit dem QIAEX II Gel Extraction Kit (Qiagen) aus der Agarose eluiert (5.2.1.5). Zum Einsatz kamen hier 15 µl QIAEX II und ein Elutionsvolumen von 30 µl Tris-HCL-Puffer (10 mM, pH 8,0) um einen DNA-Gehalt von 30 ng / µl zu erreichen. Die eluierte Plasmid-DNA kann bei -20°C gelagert werden.

5.2.1.10.2 USER-PCR

Die USER-PCR ist eine einfache Standard-PCR (5.2.1.3.1), die ihren Namen auf Grund des speziellen mit der für die USER-Technologie (Nour-Eldin et al. 2006) kompatiblen Primerdesigns trägt. Sie wird eingesetzt, um cDNAs für eine anschließende Expression rekombinanter Proteine zu amplifizieren. Die eingesetzten Primer enthalten eine spezifische Uracil reiche Sequenz. Dies erfordert den Einsatz einer speziellen Polymerase sofern *proof reading*-Aktivität gewünscht ist. (anderenfalls können gängige Polymerase verwendet werden). Im Rahmen der vorliegenden Arbeit wurde für die USER-PCR die *PfuTurbo Cx* hotstart DNA Polymerase (Agilent Technologie), eine spezielle Mutante der *Pfu* DNA Polymerase, eingesetzt.

Zusammensetzung des PCR-Ansatzes:

+	1	µl Template (20 – 50 ng)
+	0,5	µl dNTP Mix (10 mM jedes)
+	2,5	µl 10 x <i>PfuTurbo Cx</i> Reaktionspuffer
+	0,5	µl <i>Pfu</i> DNA Polymerase
+	1	µl Primer A (10 pmol / µl)
+	1	µl Primer B (10 pmol / µl)
+	18,5	µl ddH ₂ O
Σ 25		µl

PCR-Programm:

95°C	3 min	} 35 x
95°C	45 sek	
58°C	1 min	
72°C	90 sek	
72°C	10 min	
16°C	∞	

5.2.1.10.3 USER-Reaktion

Für die USER-Reaktion kommt das USER-Enzym, eine Mischung aus Uracil-DNA-Glycosylase und der DNA-Glycosylase-Lyase-Endonuclease VIII zum Einsatz. Dabei dient die Uracil-DNA-Glycosylase der Entfernung der Uracilbase bei intakt bleibendem Phosphatrückgrad der DNA. Die DNA-Glycosylase-Lyase-Endonuclease VIII generiert dann durch Schneiden der Phosphodiesterbindung ein freies 3'OH-Ende im Phosphatrückgrad. Auf Grund der speziell designten Primer entsteht so ein mit dem USER-Vektor hybridisierender Überhang, welcher auf Grund seiner Länge genug Eigenbindung erzeugt, um bei sofortiger Transformation auf eine Ligation verzichten zu können.

Eine Inkubation erfolgte für 20 min bei 37°C gefolgt von 20 min bei 25°C im PCR-Cycler.

Reaktionsansatz:

+ 8 µl Template
+ 2 µl vorbereiteter Vektor
+ 1 µl USER-Enzym
Σ 11 µl

5.2.1.10.4 USER-Transformation

Die hier aufgeführte Transformation unterscheidet sich von Punkt 5.2.1.7 wie folgt: 100 µl kompetente *E. coli*-Zellen des Stammes XL1Blue (Agilent Technologies) (5.1.2.2.1) wurden auf Eis aufgetaut und mit der Plasmid-DNA vorsichtig auf Eis gemischt. Nach 30 min Inkubation auf Eis wurde das Gemisch für 75 sek einem Hitzeschock bei 42°C im Wasserbad ausgesetzt. Anschließend folgte eine sofortige Abkühlung für 2 min auf Eis und eine Zugabe von 200 µl SOC-Medium. Nach einer Inkubation von 60 min bei 37°C wurde der Ansatz auf LB-Platten mit 25 µl Amp (100 µg / ml) ausgestrichen und über Nacht bei 37°C inkubiert.

5.2.1.11 Isolierung von Plasmid-DNA

Die in dieser Arbeit verwendeten Methoden basieren auf der Isolation mittels alkalischer Lyse durch Natriumdodecylsulfat (SDS) (Birnboim and Doly 1979). Hauptsächlich wurde die im Folgenden beschriebene Methode verwendet. Alle Angaben für verwendete Puffer sind dem Abschnitt Puffer (5.4) zu entnehmen.

- 2 ml der *E. coli*-Kultur wurden bei 5000 rpm für 5 min zentrifugiert.
- Das entstehende Pellet wurde durch kräftiges Vortexen in 300 µl Puffer 1 mit RNase A (10 µl (10 mg / ml) pro ml Puffer) resuspendiert und nach Zugabe von 200 µl Puffer 2 für max. 5 min bei Raumtemperatur inkubiert.
- Nach Zugabe von 300 µl Puffer 3 und sofortigem Invertieren des Ansatzes wurde für 15 - 20 min auf Eis inkubiert.

- Durch Zentrifugation für 10 min bei maximal möglicher rpm und Raumtemperatur wurde der Überstand in ein neues Eppendorfgefäß überführt, mit einem Volumen Chloroform (HPLC-grade) versetzt und gut gemischt.
- Der Überstand nach 10 min Zentrifugation bei max. rpm wurde mit 0,7 Volumen Isopropanol versetzt und gut gevortext.
- Nach Zentrifugation bei maximal möglicher rpm für 10 – 30 min wurde das Pellet mit 500 µl 70 % Ethanol gewaschen und erneut für 5 min bei max. rpm zentrifugiert.
- Das Pellet wurde nach Trocknung bei Raumtemperatur in 50 µl 10 mM Tris-HCl-Puffer (pH 8,0) gelöst und kann bei -20°C gelagert werden.

Nur in Fällen schwierig zu sequenzierender Sequenzen wurde auf Grund erforderlicher hoher Reinheit das Kit Wizard® Plus SV Minipreps DNA Purification Systems (Promega) (5.1.5) nach Herstellerangaben verwendet.

5.2.1.12 Restriktion von DNA

Restriktionsenzyme sind Restriktionsendonukleasen, welche spezifisch ihre 4 - 8 Basenpaare lange Bindestelle erkennen und DNA-Stränge innerhalb dieser Erkennungssequenz oder in benachbarten Regionen hydrolisieren. Für die durchgeführten analytischen Restriktionen wurde die isolierte Plasmid-DNA mit passenden Restriktionsenzymen gespalten. Zur Kontrolle von Klonierungen in den pGEM®-T-Easy Vektor (Promega) (5.1.2.1.1) wurde hierfür das Restriktionsenzym *EcoRI* (Fermentas) (5.1.4) und für Arbeiten mit den modifizierten Varianten des pET-52b(+) Vektors (Novagen / Merck) (5.1.2.1.3, 5.1.2.1.4) das Restriktionsenzym *NcoI* (Fermentas) (5.1.4) verwendet.

Restriktionsansatz *EcoRI*:

- 3 µl Plasmid-DNA
 - 3 µl 1 x *EcoRI* Puffer
 - 0,3 µl *EcoRI*
 - 23,7 µl ddH₂O
-
- Σ 30 µl
-

Restriktionsansatz *NcoI*:

- 3 µl Plasmid-DNA
 - 1 µl 1 x Tango Puffer
 - 0,5 µl *NcoI*
 - 5,5 µl ddH₂O
-
- Σ 10 µl
-

Der Restriktionsansatz wurde für zwei Stunden oder über Nacht bei 37°C inkubiert.

5.2.1.13 Photometrische Bestimmung der DNA-Konzentration

Alternativ zur Mengenbestimmung von DNA-Fragmenten mittels Gelelektrophorese kann DNA auch UV-spektrophotometrisch vermessen werden. DNA besitzt im UV-Licht bei 260 nm ein charakteristisches Extinktionsmaximum. Mit Hilfe der Beziehung OD₂₆₀ µg / ml DNA ist die Bestimmung der DNA-Konzentration am Photometer möglich. Zur Kontrolle der DNA-Reinheit kann ein Vergleich OD₂₆₀ : OD₂₈₀ durchgeführt

werden. Bei sauberer DNA wird ein Verhältnis von 2:1 erwartet. Auch RNA-Fragmente besitzen ein Extinktionsmaximum bei 260 nm, folgen aber der Beziehung $OD_{260} = 1 \equiv 40 \mu\text{g} / \text{ml}$.

$$c [\mu\text{g} / \mu\text{l}] = \frac{OD_{260} \times 50 \times VF}{1000}$$

beziehungsweise

$$c [\mu\text{g} / \mu\text{l}] = \frac{OD_{260} \times 40 \times VF}{1000}$$

VF = Verdünnungsfaktor

Bei der photometrischen Bestimmung der DNA- oder RNA-Konzentrationen wurden 5 μl Probe mit 445 μl ddH₂O vermischt und photometrisch bei OD_{260} vermessen. Der Gehalt wurde nach der oben genannten Formel ermittelt.

5.2.1.14 DNA-Sequenzierungen

Die DNA-Sequenzierungen erfolgte durch die Firmen GATC Biotech AG (Konstanz, Deutschland) oder Eurofins MWG GmbH (Ebersberg, Deutschland).

5.2.1.15 Isolierung von genomischer DNA

Die in dieser Arbeit verwendete Methode beruht auf einer Cetylhexatrimethylammoniumbromid (CTAB)-Extraktion, bei welcher störende Polysaccharide und Proteine durch Komplexbildung mit CTAB abgetrennt werden. Die Vorschrift, welche für *Nicotiana tabacum* entwickelt wurde (Saghai-Maroo et al. 1984), eignet sich auch zur DNA-Gewinnung aus anderen Pflanzenmaterialien.

- Es wurde 1 g in Stickstoff gefrorenes und gemörsertes Blattmaterial unter Vermeidung des Auftauens in ein zentrifugierbares Gefäß geeigneter Größe eingewogen.
- Das Material wurde mit 1 ml 2 x CTAB-Puffer (auf 65°C vorgewärmt) vorsichtig vermischt und unter Zugabe von weiteren 1,5 ml 1 x CTAB-Puffer (auf 65°C vorgewärmt) homogenisiert.
- Das Gemisch wurde für 1,5 std bei 65°C schüttelnd inkubiert.
- Nach fünf minütigem Abkühlen bei Raumtemperatur wurde ein Volumen Phenol / Chloroform / Isoamylalkohol (25:24:1) zugesetzt und weitere 5 min schüttelnd inkubiert.
- Zur Phasentrennung wurde die Lösung 10 min bei 2.000 rpm und 20°C zentrifugiert und die obere DNA-haltige Phase mit einer abgeschnittenen Pipettenspitze in ein neues Zentrifugationsröhrchen überführt. War der Überstand noch nicht klar, so wurde das Ausschütteln mit Phenol / Chloroform / Isoamylalkohol (25:24:1) wiederholt.
- Unter Zusatz von 50 μl RNase A (10 mg / ml) wurde die Lösung 30 min bei Raumtemperatur inkubiert.
- Das Fälln der DNA erfolgte durch die Zugabe von 0,6 Volumen Isopropanol unter leichtem Schwenken und Drehen des Zentrifugationsröhrchens bis zum Sichtbarwerden einer DNA-Flocke. Eine Fällung bei -20°C über längere Zeit ist möglich. Sollte sich anstelle der Flocke ein lockerer Niederschlag bilden, kann dieser vorsichtig abzentrifugiert werden.

- Die Flocke / das Pellet wurde in 76 % Ethanol mit 0,2 M Na-Acetat (pH 5,2) gegeben und verblieb dort für 20 min.
- Anschließend wurde die Flocke / das Pellet kurz in 76 % Ethanol mit 10 mM NH₄-Acetat (pH 5,2) überführt und im Folgenden luftgetrocknet.
- Die DNA wurde über Nacht bei 4°C in 200 – 500 µl Tris-EDTA-Puffer (TE-Puffer) (pH 8,0) gelöst. Sollte sich die DNA nicht lösen, ist ein wechselweises Einfrieren und Erhitzen auf max. 65°C möglich.
- Eine Abschätzung der DNA-Konzentration erfolgte mit Hilfe eines 1 %igen Agarosegels (5.2.1.4), das mit ca. 5 µl DNA in 10 µl ddH₂O und 2 µl Ladepuffer beladen wird. Die DNA-Ausbeute liegt bei ca. 20 - 30 µg DNA pro 1 g Ausgangsmaterial.

5.2.1.16 Southern-Kapillar-Blot

Der Southern-Blot dient dem Nachweis von spezifischen Gensequenzen in einem DNA-Gemisch mit Hilfe spezifischer und markierter DNA-Sonden (Southern 1975). Hierfür wurde zunächst genomische DNA aus den zu untersuchenden Zellen isoliert (5.2.1.15). Diente die gewonnene DNA als Template für eine PCR, so wurde das PCR-Produkt ohne weitere Vorbehandlung direkt in einem Agarosegel aufgetrennt und die DNA-Fragmente anschließend, wie unter Punkt 5.2.1.17.2 beschrieben, denaturiert und depuriniert. Sollte die genomische DNA direkt für den Southern-Kapillar-Blot eingesetzt werden, wurde die DNA-Menge photometrisch bestimmt (5.2.1.13)

5.2.1.17 Restriktionsverdau

20 µg genomischer DNA wurden mit je 100 Units der Restriktionsenzyme *Bam*HI (Hochkonzentriert) (HC) (5.1.4) oder *Eco*RI HC (Fermentas) (5.1.4) behandelt. Hierfür wurde ein Restriktionsansatz (5.2.1.12) mit einem Gesamtvolumen von 100 µl gewählt und über Nacht bei 37°C inkubiert.

5.2.1.17.1 Natriumacetat-Fällung des Restriktionsansatzes

Zum Auftrag in einem Agarosegel war es notwendig, die erhaltenen DNA-Fragmente aus dem 100 µl Restriktionsansatz durch Fällung auf ein Volumen von 20 µl zu konzentrieren. Im Rahmen dieser Arbeit wurde dazu eine Na-Acetat-Fällung durchgeführt.

- 100 µl Restriktionsansatz wurden mit 250 µl 75 mM Natriumacetat (NaAc) in 96 % Ethanol (pH 5,0) versetzt und die DNA-Fragmente durch Zentrifugation bei 20.000 g für 15 min pelletiert.
- Der Überstand wurde verworfen und das Pellet in 200 µl 70 % Ethanol gewaschen.
- Nach erneuter Zentrifugation für 6 min bei 20.000 g wurde das Pellet luftgetrocknet und in 20 µl TE aufgenommen.

5.2.1.17.2 Denaturierung und Depurinierung der DNA-Fragmente

Die DNA-Fragmente wurden nach Zusatz von 4 µl 6 x Ladepuffer mittels Agarose-Gelelektrophorese langsam (!) (40 V 30 min, 60 V 50 min, 70 V 1 std) aufgetrennt (5.2.1.4) und die Banden mit Ethidiumbromid (0,25 µg / ml) unter UV-Licht (366 nm) dokumentiert. Anschließend wurden die DNA-Fragmente im Agarosegel denaturiert und depuriniert, damit sie einzelsträngig und in kleineren Fragmenten vorlagen und sowohl eine bessere Übertragung auf eine Membran, als auch eine spezifische Bindung der markierten DNA-Sonden nach dem Blotting ermöglichten.

- Das Agarosegel wurde durch Schwenken für max. 10 min in 0,25 M HCl depuriniert und im Anschluss für 5 min schwenkend in ddH₂O gewaschen.
- Für die Denaturierung der DNA-Fragmente wurde das Gel 2 x in Denaturierpuffer gegeben und für 15 min vorsichtig bewegt.
- In ddH₂O wurde das Agarosegel anschließend für 5 min schwenkend gewaschen. Und für 2 x 15 min in Neutralisationspuffer vorsichtig bewegt. Der pH-Wert des Agarosegels sollte nun bei pH 9,0 liegen.

5.2.1.17.3 Aufbau des Kapillar-Blots

Der Begriff Kapillar-Blot bezieht sich auf die gewählte Methode der Übertragung der DNA-Fragmente auf eine Membran. Bei dem Kapillar-Verfahren wird ein Transferpuffer mittels Filterpapieren durch das Agarosegel gesaugt. Die DNA-Fragmente werden dabei aus dem Agarosegel herausgelaugt und bleiben auf der Membran hängen.

- Eine große Plastikbox (1) wurde mit 20 x „Saline-Sodium-Citrate“-Puffer (SSC-Puffer) (2) gefüllt und eine kleine Plastikbox (3) verkehrt herum hineingestellt.
- Drei Lagen in 20 x SSC-Puffer getränktes Filterpapier (4) werden blasenfrei so auf die kleine Plastikbox (3) gelegt, dass ihre Enden in die große Plastikbox (1) hängen.
- Das Agarosegel (5) wurde blasenfrei mit der Oberseite nach unten auf die drei Lagen Filterpapier (4) gelegt und blasenfrei mit einer in ddH₂O getränkten positiv geladenen Nylonmembran (Roche) (6) überdeckt.
- Zwei in 20 x SSC-Puffer getränkte Lagen und drei trockene Lagen Whatman-Filterpapiere (ROTH) (7 + 8) wurden vor einem Stapel aus zehn Handtrockenpapieren (9) auf die Whatman-Filterpapiere (7 + 8) gelegt.
- Der komplette Aufbau wurde mit einem Gewicht von 1 kg beschwert und über Nacht ruhig stehen gelassen.

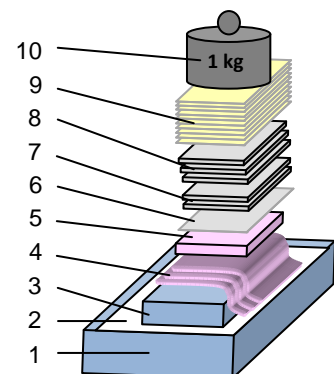


Abbildung 30: Aufbau eines Southern-Blots. Schematisch dargestellt ist die Schichtung zum Aufbau eines Southern-Blots. Weitere Details sind dem Text zu entnehmen.

5.2.1.17.4 Hybridisierung und stringentes Waschen

Nach Abbau des Kapillar-Blots wurde die Membran vorsichtig schwenkend für 5 min in 2 x SSC-Puffer gewaschen und die einzelsträngigen DNA-Fragmente durch einen Backvorgang bei 120°C für 30 min irreversibel an die Membran gebunden. Die Membran kann so behandelt trocken bei 2 – 8°C gelagert oder wie folgt direkt weiterverwendet werden:

- In auf 68°C vorgeheizter Hybridisierungslösung (ohne Sonde) wurde die Membran unter vorsichtiger Bewegung bei 68°C für 30 min prähybridisiert, um ungebundene DNA-Reste abzuwaschen.
- 25 ng / ml Sonde wurden durch fünfminütiges Kochen bei 95°C und anschließender sofortiger Kühlung auf Eis denaturiert.
- Die denaturierte Sonde wurde in auf 68°C vorgewärmte Hybridisierungslösung (ohne Sonde) gegeben und ohne Schäumen gemischt.
- Die Hybridisierungslösung ohne Sonde wurde durch die Hybridisierungslösung mit Sonde ersetzt und die Membran schwenkend bei 43°C (berechnet nach der im DIG-High Prime DNA Labeling and Detection Starter Kit I angegebenen Formel) über Nacht inkubiert.
- Die Hybridisierungslösung mit Sonde wurde abgegossen. Sie konnte für spätere Versuche aufbewahrt und wiederverwendet werden.
- Um überschüssige und unspezifische gebundene Sondenreste zu entfernen, wurde die Membran zweimal für 5 min in 0,1 % SDS / 2 x SSC-Puffer bei 15 – 25°C schwenkend gewaschen.
- Anschließend wurde die Membran bei 65°C weitere zwei Male in 0,1 % SDS / 0,5 x SSC-Puffer schwenkend gewaschen.
- Anschließend erfolgte die Äquilibration bei 15 – 25°C für 5 min schwenkend in 2 x SSC.

5.2.1.17.5 Detektion

Der verwendete Anti-digoxigenin-AP Conjugate Antikörper (Roche) ist ein Konjugat aus einem Anti-Digoxigenin-Antikörper (aus dem Schaf) und dem Enzym alkalische Phosphatase, das durch eine Farbreaktion nachgewiesen werden kann. Während der Inkubation im Dunklen wird das Substrat BCIP von der alkalischen Phosphatase dephosphoryliert und reagiert mit Sauerstoff zu einem tiefblauen Farbstoff. Bei dieser Reaktion werden Protonen freigesetzt, welche wiederum das zur Farbverstärkung eingesetzte Nitroblau-Tetrazoliumchlorid (NBT) zu dem tiefblauen Farbstoff Di-Formazan reduzieren (McGadey 1970). Da beide Farbstoffe in unmittelbarer Nähe der Sonde ausfallen und die Sonde spezifisch für das gesuchte DNA-Fragment ist, zeigt eine Färbung auf der Membran das Vorhandensein der gesuchten DNA-Fragmente auf.

- Zur Blockierung unspezifischer Bindungsstellen wurde die Membran 30 min schwenkend bei 15 - 25°C in Blockpuffer inkubiert.
- Zum Nachweis der Sonde wurde die Membran mit einem spezifisch gegen die Digoxigeninreste der Sonde gerichteten Antikörper (Anti-digoxigenin-AP Conjugate, Roche) inkubiert. Dazu wurde die Membran 30 min bei 15 – 25°C in Antikörperlösung geschwenkt.
- Um ungebundenen Antikörper abzuspielen, wurde die Membran anschließend zwei Mal für 15 min bei 15 – 25°C in Waschpuffer gegeben.

- Die Membran wurde zum Erreichen des, für die folgende Farbreaktion notwendigen pH Wertes von 9,5 für 5 min bei 15 – 25°C in Detektionspuffer geschwenkt.
- Im Dunklen wurde die Membran bis zu 16 std ruhend bei Raumtemperatur in der NBT und 5-Brom-4-Chlor-3-Indoxylphosphat (BCIP) enthaltenden Farbelösung inkubiert und anschließend in TE-Puffer (pH 8,0) gewaschen.

5.2.1.17.6 Stripping der Membran

Unter Stripping versteht man das Ablösen der spezifisch an die DNA-Fragmente gebundenen Sonde. Die DNA-Fragmente selber verbleiben hierbei auf der Membran, da sie durch den Backvorgang irreversibel verankert vorliegen. Für den Strippingvorgang wurde Dimethylformamid (DMF) unter einem Abzug auf einem Wasserbad vorsichtig bis auf 50°C - 60°C erhitzt und die Membran anschließend in das vorgewärmte DMF gegeben, wodurch sich die ausgefallenen Farbstoffe in Lösung begeben. Durch die Wärme wurden zudem die Basenpaarungen zwischen den DNA-Fragmenten und der Sonde aufgebrochen. Nach einem vorsichtigen Waschschrift in ddH₂O wurde die Membran zwei Mal bei 37°C in jeweils frische Rehybridisierungslösung gegeben. Diese basische Lösung bewirkt durch die Verwendung des alkali-labilen DIG-11-dUTP zur Sondengeneration die Dissoziation von Antigen (Sonde) und Antikörper von den DNA-Fragmenten. Nach einem abschließenden Waschschrift oder längerer Aufbewahrung in 2 x SSC-Puffer kann die Membran direkt für eine neue Prähybridisierung (5.2.1.17.4) eingesetzt werden.

5.2.1.17.7 Southern-Dot-Blot

Der Dot-Blot stellt eine Vereinfachung des eigentlichen Blot-Vorganges dar. Die zu detektierenden DNA-Fragmente müssen nicht durch eine Agarose-Gelelektrophorese aufgetrennt werden sondern befinden sich in einer Lösung, welche direkt punktförmig (englisch: dot) auf die Membran aufgebracht wird. Im Fall der vorliegenden Arbeit wurde eine Plasmid-DNA enthaltende Lösung verwendet. Die verwendete Plasmid-DNA wurde vor Verwendung auf definierte Konzentrationen von 5 pg, 15 pg und 25 pg verdünnt, um eine Einschätzung der nachweisbaren Größenordnung zu erhalten. Da Plasmid-DNA doppelsträngig vorliegt, wurde sie durch fünfminütiges Aufkochen der verdünnten Lösung bei 95°C denaturiert und sofort auf Eis abgekühlt um Doppelstrangbildung zu verhindern. Das berechnete Volumen der Lösung der Plasmid-DNA wurde dann auf die mit 2 x SSC-Puffer befeuchtete Membran aufgetragen und die Membran im Anschluss wie unter „Hybridisierung und Stringentes Waschen“ (5.2.1.17.4) beschrieben weiter behandelt.

5.2.1.17.8 Sondengeneration

Die Generation der genspezifischen DNA-Sonde erfolgte im Rahmen dieser Arbeit mittels eines PCR-Ansatzes unter Zuhilfenahme des PCR DIG Probe Synthesis Kit (Roche). Dieses Kit nutzt Uridinnukleotide, welche in C-5 Position mit Digoxigenin verlinkt wurden (DIG-11-dUTP). Im Kit sind diese Nukleotide in dem 10 x dNTP-Mix (PCR DIG Probe Synthesis Mix, Roche) in einer Mischung aus dATP, dCTP, dGTP (2 mM jedes), 1,3 mM

dTTP und 0,7 mM DIG-11-dUTP enthalten. Während einer PCR-Reaktion mit einer Taq-Polymerase werden die Digoxigenin verlinkten Uridinnukleotide statistisch in den zu amplifizierenden DNA-Strang eingebaut und können später durch einen spezifisch gegen Digoxigenin gerichteten Antikörper detektiert werden (5.2.1.17.5). Die verwendete PCR-Technik entspricht einer Standard-PCR (5.2.1.3.1) unter Einsatz genspezifischer Primer (5.1.3) und wurde auf die einzusetzenden Templates aus isolierter Plasmid-DNA (5.2.1.11) optimiert.

Zusammensetzung des PCR-Ansatzes:

– 75	pg ESP-Template, 100 pg NSP-Template
– 2,5	µl 10 x PCR DIG Probe Synthesis Mix (Vial 2)
– 2,5	µl 10 x dNTP stock solution (Vial 4)
– 5	µl 10 x PCR buffer with MgCl ₂ (Vial 3)
– 0,75	µl Enzyme mix, Expand High Fidelity (Vial 1)
– 1	µl Primer A (10 pmol / µl)
– 1	µl Primer B (10 pmol / µl)
– 36,54	µl ddH ₂ O
Σ 50	µl

PCR-Programm:

95°C	2 min	
95°C	30 sek	} 10 x
56°C / 65°C	30 sek	
72°C	40 sek	
95°C	30 sek	} 20 x
56°C / 65°C	30 sek	
72°C	40 sek	
+ 20 sek je Zyklus		
72°C	7 min	
16°C	∞	

Die angegeben Zusammensetzung des PCR-Ansatzes wurde vor der PCR-Reaktion in vier Ansätze zu je 12,5 µl aliquotiert, da in größeren Aliquotes keine Sonde generiert werden konnte. Die erfolgreiche Sondengeneration wurde durch Agarose-Gelelektrophorese (5.2.1.4) überprüft. Eine DNA-Mengenbestimmung erfolgte im Anschluss mittels Agarose-Gelelektrophorese (5.2.1.4).

5.2.2 Proteinbiochemische Methoden

5.2.2.1 Überexpression rekombinanter Proteine in *E. coli*

Für die Überexpression rekombinanter Proteine in *E. coli* wurde im Rahmen dieser Arbeit das für die USER-Technologie modifizierte pET-Expressionssystem gewählt. Der vollständige ORF der interessierenden DNA wurde mittels USER-Reaktion (5.2.1.10.3) in entsprechende Expressionsvektoren (5.1.2.1.3, 5.1.2.1.4) eingebracht und die erfolgreiche Klonierung durch Sequenzierung (5.2.1.14) bestätigt. 100 ng aufgereinigter Plasmid-DNA (5.2.1.11) wurden in den *E. coli*-Stamm BL21(DE3) pLysS (Invitrogen) (5.1.2.2.2) transformiert (5.2.1.7). Mit einer erhaltenen Einzelkolonie wurde am folgenden Tag TB-Medium (Terrific-Broth-Medium) (antibiotikahaltig) beimpft, das zwei Tage schüttelnd (230 rpm) bei 18°C inkubiert wurde. Das TB-Medium wurde mit einer Endkonzentration von 100 µg / ml Amp und 34 µg / ml Chloramphenicol (Chl) versetzt. Zusammensetzung siehe 5.3.

Ausgehend von der Vorkultur wurde nach zwei Tagen eine Hauptkultur in TB-Medium mit Antibiotika auf eine OD₆₀₀ von 0,1 – 0,15 eingestellt. Je nach gewünschter Proteinmenge betrug das Gesamtvolumen der Hauptkultur 50 ml oder 250 ml.

Wenn durch schüttelnde (230 rpm) Inkubation bei 18°C eine OD₆₀₀ zwischen 0,4 und 0,6 erreicht war, wurden 50 µl / 250 µl 1 M IPTG für eine Endkonzentration von 1 mM zugegeben und weitere 16 std bei 18°C schüttelnd (230 rpm) inkubiert.

5.2.2.2 Proteinextraktion aus *E. coli*

- Die vollständige Hauptkultur wurde in ein Zentrifugengefäß passender Größe überführt und die *E. coli*-Zellen wurden durch Zentrifugation bei 3.500 rpm und 6°C für 5 – 10 min sedimentiert.
- Der Überstand wurde verworfen, das Pellet nach Abtropfen in 1,5 ml (2-Morpholinoethansulfonsäure-Monohydrat-Puffer) MES-Puffer (50 mM, pH 6,0) pro 1 mg Pellet vorsichtig resuspendiert.
- Die Zellen wurden anschließend unter Kühlung in Eiswasser durch Sonifizierung (Branson Sonifier 250; BRANSON Ultrasonics Corporation, Connecticut, USA) für 2 x 90 sek mit 30 sek Pause aufgeschlossen.
- Zelltrümmer wurden durch Zentrifugation bei 10.000 rpm für 30 min bei 6°C pelletiert und der Überstand als Rohextrakt bis zur weiteren Verwendung gekühlt gelagert.

5.2.2.3 Aufreinigung rekombinanter Proteine aus *E. coli*-Rohextrakt

Für eine *in vitro* Aufreinigung der rekombinant exprimierten Proteine ist es notwendig, dass ein Vektor gewählt wird, der den gewünschten ORF um die Codons eines Tags ergänzt. Im Rahmen der vorliegenden Arbeit wurden rekombinante Proteine mit einem *Strep-Tag-II* versehen und mittels des Strep-tag® purification system (IBA) (5.1.5) nach Herstellerangaben aufgereinigt. Das zu wählende Säulenvolumen betrug ein Hundertstel des Volumens der eingesetzten Hauptkultur.

5.2.2.4 Kolorimetrische Bestimmung der Proteinkonzentration

Das verwendete Pierce® BCA Protein Assay Kit (Thermo Fisher Scientific) (5.1.5) kombiniert die Reduktion von Cu²⁺-Ionen in wässrig alkalischer Lösung durch Proteine zu Cu¹⁺-Kationen mit einer hoch sensitiven und selektiven kolorimetrischen Bestimmung der Cu¹⁺ durch Einsatz einer Bicinchonsäure (BCA) enthaltenden Lösung (Smith et al. 1985). Bei 562 nm detektierbare Violettffärbung der Lösung basiert auf der Komplexbildung zwischen zwei BCA-Molekülen und einem Cu¹⁺. Zur Erstellung einer Kalibrationsgeraden wurde Rinderalbumin (BSA) in einer Verdünnungsreihe bekannter Konzentrationen eingesetzt. Der Assay und die photometrische Bestimmung erfolgte in einer Mikrotiterplatte unter Verwendung des Microplate Readers Sunrise™ (Tecan Group Ltd., Männedorf, Schweiz).

- Bei Einsatz des Rohextraktes wurde dieser über zwei Schritte hundertfach verdünnt. Für eine genaue Bestimmung sollten drei verschiedene Mengen eingesetzt werden.

- Puffer A und Puffer B (Bestandteile des verwendeten Kits) wurden in ausreichender Gesamtmenge im Verhältnis 50:1 gemischt.
- Eine Verdünnungsreihe (0 mg – 1000 mg) wurde aus eine BSA-Stammlösung (2 mg / ml) erstellt. Dabei wurde derjenige Puffer verwendet, in welchem auch die Proteinextrakte vorliegen.
- In einer Mikrotiterplatte wurden je Vial 200 µl Puffergemisch vorgelegt und mit 25 µl Probe beziehungsweise BSA-Lösung versetzt.
- Im Anschluss an eine Inkubation von 30 min bei 37 °C im Microplate Reader erfolgte die Absorptionmessung bei 562 nm. Zur Berechnung wurde durch lineare Regression eine Kalibriergerade erstellt. Die Gradengleichung wurde zur Berechnung des Proteingehaltes der Proben verwendet.

5.2.2.5 SDS-Polyacrylamid-Gelelektrophorese

Bei dieser Variante der Gelelektrophorese dient Polyacrylamid als Trennungsmatrix. 10 µl einer Proteinprobelösung wurden mit 10 µl 3 x Ladepuffer (SDS haltig) und 10 µl ddH₂O versetzt und für 5 min bei 95°C inkubiert. Die SDS-Polyacrylamidgele wurden nach der Elektrophorese-Methode nach Laemmli (Laemmli 1970) hergestellt und waren 15 %ig (Tabelle 15). Eingesetzt wurden Gelkammern mit –träger, -spacern und -kämmen der Firma Bio-Rad Laboratories GmbH (München, Deutschland). Für den Gellauf wurde eine Spannung von 180V für 50 min in 1 x SDS-Laufpuffer verwendet. Die Detektion der Proteine erfolgte durch Coomassie-Brillant-Blau-Färbung und beruhte auf der Anlagerung dieses Triphenylmethanfarbstoffes an die basischen Seitenketten der Aminosäuren. Proteine wurden dadurch unspezifisch angefärbt und konnten optisch detektiert werden. Anhand des mitgeführten Standards bekannter Größen (Fast-Ruler DNA Ladder, Fermentas), konnte eine Größenzuordnung erfolgen.

Tabelle 15: Pipettierschema für zwei Polyacrylamidgele (15 %).

Schema für zwei Gele	Sammelgel	Trenngel
Sammelgel-Puffer	2,5 ml	-
Trenngel-Puffer	-	5,0 ml
Acrylamid	1,5 ml	10 ml
ddH ₂ O	6,0 ml	4,7 ml
10 % SDS	100µl	200 µl
10 % Ammoniumpersulfat (APS)	30µl	100 µl
N, N, N', N'-Tetramethylethylenediamine (TEMED)	10µl	10 µl

5.2.2.6 Assay spezifizierender Proteine

Myrosinase und Glucosinolate sind in getrennten Zellkompartimenten gelagert. Kommen sie durch Zerstörung der Zelle zusammen, so hydrolisiert die Myrosinase das Glucosinolat und es entsteht ein instabiles Aglycon. Dieses Aglycon dient den zu charakterisierenden

spezifizierenden Proteinen als putatives Substrat. Durch die Instabilität des Aglycons ist es nicht möglich, dieses als Substrat direkt zuzuführen. Es wurde daher ein Assayansatz gewählt in welchem das Glucosinolat und die Myrosinase enthalten waren, so dass ein Umsatz zum Aglycon das Substrat für die spezifizierenden Proteine lieferte. Die erhaltenen Glucosinolat-Hydrolyseprodukte wurden detektiert. An Hand ihres Profils war eine Einordnung der spezifizierenden Proteine als NSP, ESP oder TFP möglich. In einem Gesamtvolumen von 500 µl MES-Puffer (50 mM, pH 6,0) wurden 1,5 mg Gesamtprotein als bakterieller Rohextrakt, 1 mM Benzylglucosinolat oder 2 mM Allylglucosinolat und 0,005 Units Myrosinase inkubiert. Ein typisches Pipettierschema ist in Tabelle 16 angegeben.

Tabelle 16: Pipettierschema für den Assay spezifizierender Proteine.

	5 µl → 100 mM Benzylglucosinolat,	10 µl → 100 mM Allylglucosinolat
H ₂ O	125 ml	120 ml
0,5 M MES, pH 6,0	20 ml	20 ml
Rohextrakt (in 50 mM MES)	x ml	x ml
Myrosinase (0,005 Units)	50 µl	50 µl
50 mM MES	auf 500 µl	auf 500 µl

Der Ansatz wurde für 30 min bei Raumtemperatur inkubiert, mit 50 µl Phenylcyanid (1:10000) als internem Standard versetzt und zweimal mit je circa 750 µl Dichlormethan ausgeschüttelt. Die erhaltenen Dichlormethanphasen wurden vereinigt und über ein Natriumsulfat-Säulchen von Wasser befreit. Mitgeführt wurden jeweils eine Negativkontrolle ohne Zusatz von Rohextrakt sowie eine Negativkontrolle der statt des Rohextraktes mit rekombiniertem Protein ein Rohextrakt aus *E. coli* zugesetzt wurde, die nur den „leeren“ Vektor (ohne Insert) enthielt. Die Detektion der Hydrolyseprodukte erfolgte nach Einengen der wasserfreien Dichlormethanphasen unter einem Luftstrom auf ein Volumen von 100 µl - 200 µl mittels GC-MS (5.2.3.4).

5.2.3 Chemisch-analytische Methoden

5.2.3.1 Isolierung von Desulfoglucosinolaten

Die Aufarbeitung von Glucosinolaten in Form ihrer desulfatierten Analoga, den Desulfoglucosinolaten, diente im Rahmen dieser Arbeit der Analyse des Glucosinolatprofils von in stickstoffgefrorenem und gefriergetrocknetem Pflanzenmaterial. Durch Zugabe eines internen Standards konnten die identifizierten Glucosinolate zudem quantifiziert werden. Zur Aufreinigung der Desulfoglucosinolate wurde der Anionenaustauschers Sephadex in Verbindung mit dem Sulfatester-hydrolysierenden Enzym Sulfatase verwendet (Hogge 1988). Die Durchführung erfolgte entsprechend (Burow et al. 2006b). Abweichend wurde der dritte

Waschschritt der DEAE-Sephadex-Säule mit 1 ml 0,02 M Natriumacetat-Puffer durchgeführt. Eingesetzt wurden 100 mg stickstoffgefrorenes Pflanzenmaterial.

5.2.3.2 Hochleistungsflüssigkeitschromatographie (HPLC) – Analyse von Desulfoglucosinolaten

Bei der HPLC handelt es sich um ein Trennverfahren, bei welchem die Probenflüssigkeit mit Hilfe einer flüssigen Phase (Eluent) über eine stationäre Phase (Trennsäule) gepumpt wird. Dies geschieht unter hohem Druck. Im Rahmen dieser Arbeit wurde einem HPLC 1200 Series-Gerät (Agilent) das mit einer LiChrospher RP18-Säule (250 x 4.6 mm, 5 µm Partikelgröße (Wicom)) und einem DAD-Detektor ausgestattet war. Die Detektion erfolgte bei 229 nm ($\lambda = 190 - 360$ nm; 2,00 nm Schritte). Substanzen wurden an Hand ihrer Retentionszeiten und UV-Absorptionsspektren im Vergleich mit Standards (Reichelt et al. 2002) identifiziert. Die Quantifizierung erfolgte über einen Vergleich der Peakflächen (229 nm) mit der Peakfläche des internen Standards unter Verwendung experimentell ermittelter Responsfaktoren (2,0 Aliphatische Glucosinolate, 0,5 Indolische Glucosinolate, 1,8 Benzylglucosinolat) (Burow et al. 2006b; Sarosh et al. 2010). Das Standard Injektionsvolumen betrug 10 µl, die Temperatur wurde konstant auf 23°C gehalten. Die Flussrate des Fließmittels betrug 1,0 ml / 1 min mit dem im Folgenden angegebenen Gradienten.

Gradientenstufen	1	2	3	4	5	6
Zeit (min)	0	1	6	8	18	23
% Eluent A (Wasser)	98,5	98,5	95,0	93,0	79,0	71,0
% Eluent B (Acetonitril)	1,5	1,5	5,0	7,0	21,0	29,0

Gradientenstufen	7	8	9	10	11
Zeit (min)	30	30,5	34,5	35	42
% Eluent A (Wasser)	57,0	7,0	7,0	98,5	98,5
% Eluent B (Acetonitril)	43,0	93,0	93,0	1,5	1,5

5.2.3.3 Analyse von Glucosinolat-Hydrolyseprodukten

Zur Analyse von Glucosinolat-Hydrolyseprodukten in Pflanzenmaterial wurde frisches pflanzliches Gewebe homogenisiert und nach Zusatz eines internen Standards mittels GC-MS analysiert. Im Einzelnen wurde entsprechend (Burow et al. 2006b) vorgegangen. Eingesetzt wurden 100 mg frisches Pflanzenmaterial. Als Standard wurden 50 µl Phenylcyanide (1 : 10000 in Methanol) zugegeben. Zur Analyse der Glucosinolat-Hydrolyseprodukte wurden diese mit Dichlormethan ausgeschüttelt und unter einem Luftstrom bis auf ein Volumen von 100 µl - 200 µl eingeengt und zur Vermessung mittels GC-MS verwendet.

5.2.3.4 Gaschromatographische (GC) - Analyse von Hydrolyseprodukten

Bei der Gaschromatographie (GC) handelt es sich um ein Trennverfahren, bei welchem die Probenflüssigkeit mit Hilfe eines, als mobile Phase dienenden, Gasstromes (Trärgas) über eine stationäre Phase (Trennsäule) geleitet wird. Dies geschieht unter hohen Temperaturen. Im Rahmen dieser Arbeit wurde ein Gaschromatograph (Apparatur GC 6890N Series (Agilent)) verwendet, der mit einer Kapillarsäule (HP5MS, 30 m x 0.25 mm x 0.25 μ m (Wicom)) ausgestattet und zur Detektion mit einem Massenspektrometer (5975B inert XL EI / CI MSD (Agilent)) gekoppelt war (GC-MS). Als Trärgas wurde Helium verwendet. Zur Identifizierung der Hydrolyseprodukte wurden die Retentionszeiten und Massenspektren bekannter Standards oder bereits veröffentlichte Massenspektren herangezogen (Spencer and Daxenbichler 1980). Das Injektionsvolumen der Probe betrug 1 μ l bei 200°C. Weitere Details des genutzten Temperaturprogrammes und den Bedingungen der Elektronenstoßionisation sind im Folgenden angegebenen.

Temperaturprogramm				
Temperatur (°C)	35	12 min ⁻¹ - 280	30 min ⁻¹ - 300	300

Elektronenstoßionisation	
Grenzflächentemperatur	280°C
Repeller	30 V
Emission	34.6 μ A
Elektronenenergie	70 eV
Inonenquellentemperatur	230°C

5.2.4 Bioinformatische Methoden

5.2.4.1 Computergestützte Sequenzanalysen

5.2.4.1.1 Analyse der Domänenstruktur von Aminosäuresequenzen

Zur Analyse der Domänenstruktur wurde die InterProScan Sequence Search Option des EMBL-EBI Webservers genutzt ((Zdobnov and Apweiler 2001), <http://www.ebi.ac.uk/Tools/pfa/iprscan/>). InterProScan bietet eine Zusammenfassung mehrerer unabhängiger Datenbanken zur *in silico*-Identifizierung von Domänen oder funktionellen Gruppen innerhalb einer gegebenen Aminosäuresequenz. Im Rahmen der vorliegenden Arbeit wurde für derartige Strukturanalysen ausschließlich die Option der Pfam Proteindomänen-Familien Datenbank genutzt (Finn et al. 2010).

5.2.4.1.2 Erstellung multipler Sequenzalignments von Aminosäuresequenzen

Multiple Sequenzalignments von Aminosäuresequenzen wurden unter Verwendung des EBI Webservers für ClustalW2 ((Chenna et al. 2003), <http://www.ebi.ac.uk/Tools/msa/clustalw2/>)) erstellt. Hierbei wurden die Standardeinstellungen des Programmes für Proteine ohne Nummerierung des Alignments verwendet. Als Eingabeformat wurde das FASTA-Format einzelner Aminosäuresequenzen gewählt. Ein weiterer zur Verfügung stehender Webserver zur Erstellung multipler Sequenzalignments namens MultAlin ((Corpet 1988), <http://multalin.toulouse.inra.fr/multalin/>)) wurde ebenfalls unter Einsatz des Eingabeformates FASTA-Format und der Standardeinstellung verwendet.

Für die graphische Darstellung wurde auf dem EMBnet Webservers BOXSHADE 3.21 (geschrieben von K. Hofmann und M. Baron, http://www.ch.embnet.org/software/BOX_form.html) zurückgegriffen. An Hand ihrer biochemischen Eigenschaften ordnet das Programm Aminosäuren in neun Gruppen (FYW / ILVM / P / DE / GA / ST / NQ / C / RKH). Hierbei sind Aminosäuren aus einer Gruppe als ähnlich definiert. Schattierungen wurden eingefügt, wenn mindestens 80 % der Aminosäuren in einer Position identisch (schwarzer Grund, weiße Schrift) oder ähnlich (grauer Grund, weiße Schrift) waren.

5.2.4.2 Phylogenetische Analysen

Für phylogenetische Analysen kam das Programm MEGA 5.05 (Tamura et al. 2011) zum Einsatz. Mit diesem Programm wurden phylogenetische Berechnungen von Distanzen, Selektionsdruck und phylogenetischen Bäumen vorgenommen. Außerdem wurden die für die Analysen genutzten Alignments sowohl auf Ebene der Nukleotidsequenzen als auch auf Ebene der Aminosäuresequenzen erstellt und nachträglich per Hand nach der Pam250-Matrix (Dayhoff et al. 1978) editiert. Weitere Details der genutzten Einstellungen sind im Folgenden angegeben.

Alignment		ClustalW (Thompson et al. 1994) und Muscle (Edgar 2004) unter Standardeinstellungen.
Substitutionsmodell	Nukleotide	Kimura-2-Parameter-Modell (Kimura 1980) korrigiert für invariable Positionen (+I) und für positionsspezifisch variable Raten (+G) mit 0,8.
	Aminosäuren	WAG-Modell (Whelan and Goldman 2001) korrigiert für positionsspezifisch variable Raten (+G) mit 0,8.
	Codons	Equal-input Modell korrigiert für positionsspezifisch variable Raten (+G) mit 0,8.
Methoden zur Stammbaumkonstruktion	Neighbor Joining-Verfahren (Saitou and Nei 1987)	Einbezogen wurden Transitionen und Transversionen sowie alle Codonpositionen. Alignmentlücken wurden vernachlässigt, wenn mehr als 95 % der Sequenzen in einer Position uninformativ waren.
	Maximum Likelihood-Methode (Fisher 1912; Aldrich 1997)	Alle Codonpositionen wurden einbezogen. Alignmentlücken wurden vernachlässigt, wenn mehr als 95 % der Sequenzen in einer Position uninformativ waren. Der Ausgangsbaum wurde automatisch erstellt. Als heuristische Methode wurde „Close-Neighbor-Interchange,“ verwendet.
	Maximum-Parsimonie-Methode (Camin 1965)	Alle Positionen die eine Alignment Lücke beinhalteten, wurden vernachlässigt. Als heuristische Methode wurde „Close-Neighbor-Interchange,“ mit einem einfachen Suchlevel für die Ausgangsstammbäume unter zufälliger Addition von Sequenzen in zehn Wiederholungen verwendet.

5.3 Medien

Im Rahmen der vorliegenden Arbeit wurden überwiegend Chemikalien verwendet, die von den Firmen Merck KGaA (Darmstadt, Deutschland), Sigma-Aldrich Chemie GmbH (Steinheim, Deutschland), und CARL ROTH GmbH + Co. KG (Karlsruhe, Deutschland) bezogen wurden.

<u>LB-Medium (Miller):</u>	10 g	Pepton aus Casein
Lysogeny-Broth-Medium /	5 g	Hefeextrakt
Luria-Bertani-Medium	10 g	Natriumchlorid
	auf 1 l mit ddH ₂ O auffüllen, pH 7,0 – 7,2 einstellen, autoklavieren	
 Festmedium:	15 g	Agar agar
	autoklavieren	
 <u>LB-Glycerin-Medium:</u>	1 g	Pepton aus Casein
	0,5 g	Hefeextrakt
	1 g	Natriumchlorid
	auf 40 ml mit ddH ₂ O auffüllen, pH 7,5 einstellen	
	mit 60 ml Glycerin auf 100 ml auffüllen, autoklavieren	
 <u>SOC-Medium:</u>	2.0 g	Pepton aus Casein
Super Optimal Broth-Medium	0.5 g	Hefeextrakt
	1 ml	1 M Natriumchlorid
	0.25 ml	1 M Kaliumchlorid
	auf 100 ml ddH ₂ O auffüllen, autoklavieren	
	1 ml	2 M Magnesium (sterilfiltriert)
	1 ml	2 M Glucose (sterilfiltriert)
	zum abgekühlten Medium zugeben.	
 <u>TB-Medium:</u>	1,2 %	Pepton aus Casein
Terrific-Broth-Medium	2,4 %	Hefeextrakt
	0,4 %	Glycerol
	0,17 M	Kaliumdihydrogenphosphat (KH ₂ PO ₄)
	0,72 M	Dikaliumhydrogenphosphat (K ₂ HPO ₄)

5.4 Puffer und Lösungen

TE-Puffer:

Tris-EDTA-Puffer

10 mM Tris-HCl (pH 8,0)
1 mM EDTA
auf 1 l mit ddH₂O auffüllen, pH 8,0 einstellen,
autoklavieren

10 x TBE-Puffer:

Tris-Borat-EDTA-Puffer

890 mM Tris Base
890 mM Borsäure
20mM EDTA
auf 1 l mit ddH₂O auffüllen, pH 8,2 – 8,4 einstellen

6 x Ladepuffer (Agarosegele):

0,025 % Xylencyanol
0,025 % Bromphenolblau
25 % Ficoll 400
50 mM EDTA
auf 100 ml mit ddH₂O auffüllen, Lösung filtrieren

Puffer 1:

1,515 g Tris-Base
0,930 g Na₂EDTA x 2 H₂O
auf 250 ml mit ddH₂O auffüllen, pH 8,0 einstellen,
autoklavieren

Puffer 2:

2 g NaOH-Pellets
in 225 ml ddH₂O lösen
Zugabe von 25 ml einer 10 %igen SDS-Lösung,
autoklavieren

Puffer 3:

73,625 g Kaliumacetat
auf 250 ml mit ddH₂O auffüllen, pH 5,5 einstellen
→ Achtung circa 27 ml Eisessig notwendig

2 x CTAB-Puffer:

Cetylhexatrimethylammoniumbromid-Puffer

1,0 g CTAB
10 ml 1 M Tris-HCl pH 8,0
2,0 ml 0,5 M EDTA pH 8,0
14 ml 5 M NaCl
auf 50 ml mit ddH₂O auffüllen, pH 7,5 einstellen,
autoklavieren
→ kurz vor Gebrauch 2 % (v / v) β-Mercaptoethanol
zugabe

Denaturierungspuffer:

1,5 M NaCl
0,5 M NaOH
auf 1 l mit ddH₂O auffüllen, autoklavieren

<u>Neutralisationspuffer:</u>	1,5 M	NaCl	
	0,5 M	Tris-Base	
	0,001 M	EDTA	
	auf 1 l mit ddH ₂ O auffüllen, pH 7,2 einstellen, autoklavieren		
<u>20 x SSC-Puffer:</u>	3 M	NaCl	
"Saline-Sodium-Citrate"-Puffer	0,3 M	Na ₃ -Citrat x 2H ₂ O	
	auf 1 l mit ddH ₂ O auffüllen, pH 7,0 einstellen, autoklavieren		
<u>Hybridisierungslösung:</u>	64 ml	ddH ₂ O	
	in zwei Schritten zu den DIG Easy Hyb Granules (DIG-High Prime DNA Labeling and Detection Starter Kit I, Roche) geben und sofort für 5 min bei 37°C rühren		
<u>Waschpuffer:</u>	0,1 M	Maleinsäure	
	0,15 M	NaCl	
	auf 1 l mit ddH ₂ O auffüllen, pH 7,5 einstellen, autoklavieren		
	0,3 % (v / v)	Tween 20 (sterilfiltriert)	zugabe
<u>Maleinsäurepuffer:</u>	0,1 M	Maleinsäure	
	0,15 M	NaCl	
	auf 1 l mit ddH ₂ O auffüllen, pH 7,5 einstellen, autoklavieren		
<u>Blockpuffer:</u>	10 x Blockpuffer (mitgeliefert im DIG-High Prime DNA Labeling and Detection Starter Kit I, Roche) mit Maleinsäurepuffer 1 : 10 auf 1 x Blockpuffer verdünnen		
<u>Antikörperlösung:</u>	Anti-digoxigenin-AP (DIG-High Prime DNA Labeling and Detection Starter Kit I, Roche) für 5 min bei 10.000 rpm zentrifugieren. Benötigtes Volumen vorsichtig von der Oberfläche abnehmen. 1 : 5000 im Blockpuffer verdünnen		
<u>Detektionspuffer:</u>	0,1 M	Tris-HCl-Puffer	
	0,1 M	NaCl	
	auf 1 l mit ddH ₂ O auffüllen, pH 9,5 einstellen, autoklavieren		
<u>Färbelösung:</u>	40 µl	NBT / BCIP (DIG-High Prime DNA Labeling and Detection Starter Kit I, Roche)	in je 2 ml Detektionspuffer geben

<u>Rehybridisierungslösung:</u>	0,2 M	NaOH
	auf 1 l mit ddH ₂ O auffüllen, autoklavieren	
	0,1 % (w / v)	SDS
<u>Sammelgel-Puffer:</u>	60,5 g	Tris-Base
	auf 1 l mit ddH ₂ O auffüllen, pH 6,8 einstellen	
<u>Trenngel-Puffer:</u>	181,5 g	Tris-Base
	auf 1 l mit ddH ₂ O auffüllen, pH 8,8 einstellen	
<u>3 x Ladepuffer (Polyacrylamidgele):</u>	4,4 ml	ddH ₂ O
	6,25 ml	Sammelgelpuffer
	5 ml	Glycerol
	1 g	SDS
	75 mg	DTT
	21 mg	Bromphenolblau
<u>10 x SDS-Laufpuffer:</u>	30,3 g	Tris-Base
Natriumlaurylsulfat-Puffer	144 g	Glycin
	10 g	SDS
	auf 1 l mit ddH ₂ O auffüllen, autoklavieren	
<u>Coomassie-Färbelösung:</u>	500 ml	Methanol
	0,5 g	Coomassie Brilliant Blau R250
	100 ml	Essigsäure
	auf 1 l mit ddH ₂ O auffüllen, Lösung filtrieren	
<u>Coomassie-Entfärberlösung:</u>	50 ml	Essigsäure
	165 ml	Ethanol
	auf 1 l mit ddH ₂ O auffüllen	

6. LITERATURVERZEICHNIS

- Adams, J., Kelso, R. and Cooley, L. (2000). "The kelch repeat superfamily of proteins: propellers of cell function." Trends Cell Biol **10**(1): 17-24.
- Al-Shehbaz, I., Beilstein, M. and Kellogg, E. (2006). "Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview." Plant Systematics and Evolution **259**(2): 89-120.
- Al-Shehbaz, I. A. and Al-Shammary, K. I. (1987). "Distribution and chemotaxonomic significance of glucosinolates in certain Middle-Eastern cruciferae." Biochemical Systematics and Ecology **15**(5): 559-569.
- Aldrich, J. (1997). "R.A. Fisher and the making of maximum likelihood " Statistical Science **12**(3): 162-176.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-3402.
- An, Y. Q., McDowell, J. M., Huang, S., McKinney, E. C., Chambliss, S. and Meagher, R. B. (1996). "Strong, constitutive expression of the Arabidopsis ACT2/ACT8 actin subclass in vegetative tissues." Plant J **10**(1): 107-121.
- Andreasson, E., Bolt Jorgensen, L., Hoglund, A. S., Rask, L. and Meijer, J. (2001). "Different myrosinase and idioblast distribution in Arabidopsis and Brassica napus." Plant Physiol **127**(4): 1750-1763.
- Bak, S., Nielsen, H. L. and Halkier, B. A. (1998). "The presence of CYP79 homologues in glucosinolate-producing plants shows evolutionary conservation of the enzymes in the conversion of amino acid to aldoxime in the biosynthesis of cyanogenic glucosides and glucosinolates." Plant Mol Biol **38**(5): 725-734.
- Bak, S., Paquette, S., Morant, M., Morant, A., Saito, S., Bjarnholt, N., Zagrobelny, M., Jørgensen, K., Osmani, S., Simonsen, H., Pérez, R., van Heeswijk, T., Jørgensen, B. and Møller, B. (2006). "Cyanogenic glycosides: a case study for evolution and application of cytochromes P450." Phytochemistry Reviews **5**(2): 309-329.
- Bartel, B. and Fink, G. R. (1994). "Differential regulation of an auxin-producing nitrilase gene family in Arabidopsis thaliana." Proc Natl Acad Sci U S A **91**(14): 6649-6653.
- Barth, C. and Jander, G. (2006). "Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense." Plant J **46**(4): 549-562.
- Bartling, D., Seedorf, M., Mithofer, A. and Weiler, E. W. (1992). "Cloning and expression of an Arabidopsis nitrilase which can convert indole-3-acetonitrile to the plant hormone, indole-3-acetic acid." Eur J Biochem **205**(1): 417-424.
- Bartling, D., Seedorf, M., Schmidt, R. C. and Weiler, E. W. (1994). "Molecular characterization of two cloned nitrilases from Arabidopsis thaliana: key enzymes in biosynthesis of the plant hormone indole-3-acetic acid." Proc Natl Acad Sci U S A **91**(13): 6021-6025.

- Beilstein, M. A., Al-Shehbaz, I. A. and Kellogg, E. A. (2006). "Brassicaceae phylogeny and trichome evolution." Am J Bot **93**(4): 607-619.
- Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R. and Mathews, S. (2010). "Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*." Proc Natl Acad Sci U S A **107**(43): 18724-18728.
- Benderoth, M., Pfalz, M. and Kroymann, J. (2009). "Methylthioalkylmalate synthases: genetics, ecology and evolution." Phytochemistry Reviews **8**(1): 255-268.
- Benderoth, M., Textor, S., Windsor, A. J., Mitchell-Olds, T., Gershenzon, J. and Kroymann, J. (2006). "Positive selection driving diversification in plant secondary metabolism." Proc Natl Acad Sci U S A **103**(24): 9118-9123.
- Benn, M. (1977). "Glucosinolates." Pure Appl Chem **49**: 197 - 210.
- Bertani, G. (1951). "Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*." J Bacteriol **62**(3): 293-300.
- Birnboim, H. C. and Doly, J. (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucleic Acids Res **7**(6): 1513-1523.
- Bjorkman, R. and Lonnerdal, B. (1973). "Studies on myrosinases. 3. Enzymatic properties of myrosinases from *Sinapis alba* and *Brassica napus* seeds." Biochim Biophys Acta **327**(1): 121-131.
- Braby, M. F. and Trueman, J. W. H. (2006). "Evolution of larval host plant associations and adaptive radiation in pierid butterflies." Journal of Evolutionary Biology **19**(5): 1677-1690.
- Brown, P. D., Tokuhisa, J. G., Reichelt, M. and Gershenzon, J. (2003). "Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*." Phytochemistry **62**(3): 471-481.
- Brzobohaty, B., Moore, I., Kristoffersen, P., Bako, L., Campos, N., Schell, J. and Palme, K. (1993). "Release of active cytokinin by a beta-glucosidase localized to the maize root meristem." Science **262**(5136): 1051-1054.
- Buckhout, T. J., Yang, T. J. and Schmidt, W. (2009). "Early iron-deficiency-induced transcriptional changes in *Arabidopsis* roots as revealed by microarray analyses." BMC Genomics **10**: 147.
- Burow, M., Bergner, A., Gershenzon, J. and Wittstock, U. (2007a). "Glucosinolate hydrolysis in *Lepidium sativum*--identification of the thiocyanate-forming protein." Plant Mol Biol **63**(1): 49-61.
- Burow, M., Losansky, A., Müller, R., Plock, A., Kliebenstein, D. J. and Wittstock, U. (2009). "The genetic basis of constitutive and herbivore-induced ESP-independent nitrile formation in *Arabidopsis*." Plant Physiol **149**(1): 561-574.
- Burow, M., Markert, J., Gershenzon, J. and Wittstock, U. (2006a). "Comparative biochemical characterization of nitrile-forming proteins from plants and insects that alter myrosinase-catalysed hydrolysis of glucosinolates." FEBS J **273**(11): 2432-2446.
- Burow, M., Muller, R., Gershenzon, J. and Wittstock, U. (2006b). "Altered glucosinolate hydrolysis in genetically engineered *Arabidopsis thaliana* and its influence on the larval development of *Spodoptera littoralis*." J Chem Ecol **32**(11): 2333-2349.

- Burow, M., Rice, M., Hause, B., Gershenzon, J. and Wittstock, U. (2007b). "Cell- and tissue-specific localization and regulation of the epithiospecifier protein in *Arabidopsis thaliana*." Plant Mol Biol **64**(1-2): 173-185.
- Bussy, A. (1840). "Sur la formation de l'huile essentielle de moutarde." J Pharm **27**: 464–471.
- Camin, J. H., and Sokal, R. R. (1965). "A method for deducing branching sequences in phylogeny." Evolution **Vol. 19**(No. 3 (Sep., 1965)): 311-326
- Case-Green, S. C. and Southern, E. M. (1994). "Studies on the base pairing properties of deoxyinosine by solid phase hybridisation to oligonucleotides." Nucleic Acids Res **22**(2): 131-136.
- Charron, C. S., Saxton, A. M. and Sams, C. E. (2005). "Relationship of climate and genotype to seasonal variation in the glucosinolate–myrosinase system. I. Glucosinolate content in ten cultivars of *Brassica oleracea* grown in fall and spring seasons." Journal of the Science of Food and Agriculture **85**(4): 671-681.
- Charron, C. S., Saxton, A. M. and Sams, C. E. (2005). "Relationship of climate and genotype to seasonal variation in the glucosinolate–myrosinase system. II. Myrosinase activity in ten cultivars of *Brassica oleracea* grown in fall and spring seasons." Journal of the Science of Food and Agriculture **85**(4): 682-690.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003). "Multiple sequence alignment with the Clustal series of programs." Nucleic Acids Res **31**(13): 3497-3500.
- Chomczynski, P. and Sacchi, N. (1987). "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction." Anal Biochem **162**(1): 156-159.
- Colangelo, E. P. and Gueriot, M. L. (2004). "The essential basic helix-loop-helix protein FIT1 is required for the iron deficiency response." Plant Cell **16**(12): 3400-3412.
- Cordingley, M. G., Register, R. B., Callahan, P. L., Garsky, V. M. and Colonno, R. J. (1989). "Cleavage of small peptides in vitro by human rhinovirus 14 3C protease expressed in *Escherichia coli*." J Virol **63**(12): 5037-5045.
- Corpet, F. (1988). "Multiple sequence alignment with hierarchical clustering." Nucleic Acids Res **16**(22): 10881-10890.
- Dayhoff, M. O., Schwartz, R. M. and Orcutt, B. C. (1978). "A model of evolutionary change in proteins." Atlas of Protein Sequence and Structure **5**(3): 345–352.
- de Kraker, J. W. and Gershenzon, J. (2011). "From amino acid to glucosinolate biosynthesis: protein sequence changes in the evolution of methylthioalkylmalate synthase in *Arabidopsis*." Plant Cell **23**(1): 38-53.
- Dicke, M. and Baldwin, I. T. (2010). "The evolutionary context for herbivore-induced plant volatiles: beyond the "cry for help". " Trends in plant science **15**(3): 167-175.
- Durham, P. L. and Poulton, J. E. (1990). "Enzymic properties of purified myrosinase from *Lepidium sativum* seedlings." Zeitschrift fuer Naturforschung, Section C, Biosciences **45**: 173-178.
- Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res **32**(5): 1792-1797.

- Efron, B. (1979). "Bootstrap Methods: Another Look at the Jackknife." The annals of Statistics **7**(1): 1-26.
- Ehrlich, P. R. and Raven, P. H. (1964). "Butterflies and Plants: A Study in Coevolution." Evolution **18**(4): 586-608
- Engler, M. J. a. R., C.C. (1982). The Enzymes. P. D. Boyer. Academic Press, New York, NY.
- Eriksson, S., Andreasson, E., Ekbom, B., Graner, G., Pontoppidan, B., Taipalensuu, J., Zhang, J., Rask, L. and Meijer, J. (2002). "Complex formation of myrosinase isoenzymes in oilseed rape seeds are dependent on the presence of myrosinase-binding proteins." Plant Physiol **129**(4): 1592-1599.
- Ettlinger, M. G., Dateo, G. P., Jr., Harrison, B. W., Mabry, T. J. and Thompson, C. P. (1961). "Vitamin C as a coenzyme: the hydrolysis of mustard oil glucosides." Proc Natl Acad Sci U S A **47**: 1875-1880.
- Fahey, J. W., Zalcmann, A. T. and Talalay, P. (2001). "The chemical diversity and distribution of glucosinolates and isothiocyanates among plants." Phytochemistry **56**(1): 5-51.
- Falk, A., Taipalensuu, J., Ek, B., Lenman, M. and Rask, L. (1995). "Characterization of rapeseed myrosinase-binding protein." Planta **195**(3): 387-395.
- Falk, K. L., Tokuhisa, J. G. and Gershenzon, J. (2007). "The effect of sulfur nutrition on plant glucosinolate content: physiology and molecular mechanisms." Plant Biol (Stuttg) **9**(5): 573-581.
- Finn, R. D., Mistry, J., Tate, J., Coghill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R. and Bateman, A. (2010). "The Pfam protein families database." Nucleic Acids Res **38**(Database issue): D211-222.
- Fisher, R. A. (1912). "On an absolute criterion for fitting frequency curves." Messenger of Mathematics **41**: 155]160.
- Franzke, A., Lysak, M. A., Al-Shehbaz, I. A., Koch, M. A. and Mummenhoff, K. (2011). "Cabbage family affairs: the evolutionary history of Brassicaceae." Trends Plant Sci **16**(2): 108-116.
- Frohman, M. A., Dush, M. K. and Martin, G. R. (1988). "Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer." Proc Natl Acad Sci U S A **85**(23): 8998-9002.
- Geshi, N. and Brandt, A. (1998). "Two jasmonate-inducible myrosinase-binding proteins from Brassica napus L. seedlings with homology to jacalin." Planta **204**(3): 295-304.
- Geu-Flores, F., Moldrup, M. E., Bottcher, C., Olsen, C. E., Scheel, D. and Halkier, B. A. (2011). "Cytosolic gamma-glutamyl peptidases process glutathione conjugates in the biosynthesis of glucosinolates and camalexin in arabidopsis." Plant Cell **23**(6): 2456-2469.
- Gigolashvili, T., Yatusevich, R., Rollwitz, I., Humphry, M., Gershenzon, J. and Flugge, U. I. (2009). "The plastidic bile acid transporter 5 is required for the biosynthesis of methionine-derived glucosinolates in Arabidopsis thaliana." Plant Cell **21**(6): 1813-1829.

- Gil, V. and MacLeod, A. J. (1980). "Studies on glucosinolate degradation in *Lepidium sativum* seed extracts." Phytochemistry **19**(7): 1369-1374.
- Gmelin, R., Virtanen, A. I. (1959). "A New Type of Enzymatic Cleavage of Mustard Oil Glucosides. Formation of Allylthiocyanate in *Thlaspi arvense* L. and Benzylthiocyanate in *Lepidium ruderales* L. and *Lepidium Sativum* L." Acta Chemica Scandinavica **13**: 1474-1475.
- Halkier, B. A. and Gershenzon, J. (2006). "Biology and biochemistry of glucosinolates." Annu Rev Plant Biol **57**: 303-333.
- Hansen, B. G., Kerwin, R. E., Ober, J. A., Lambrix, V. M., Mitchell-Olds, T., Gershenzon, J., Halkier, B. A. and Kliebenstein, D. J. (2008). "A novel 2-oxoacid-dependent dioxygenase involved in the formation of the goiterogenic 2-hydroxybut-3-enyl glucosinolate and generalist insect resistance in *Arabidopsis*." Plant Physiol **148**(4): 2096-2108.
- Hansen, B. G., Kliebenstein, D. J. and Halkier, B. A. (2007). "Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in *Arabidopsis*." Plant J **50**(5): 902-910.
- Hansen, C. H., Du, L., Naur, P., Olsen, C. E., Axelsen, K. B., Hick, A. J., Pickett, J. A. and Halkier, B. A. (2001). "CYP83b1 is the oxime-metabolizing enzyme in the glucosinolate pathway in *Arabidopsis*." J Biol Chem **276**(27): 24790-24796.
- Hartmann, T. (2007). "From waste products to ecochemicals: Fifty years research of plant secondary metabolism." Phytochemistry **68**(22-24): 2831-2846.
- Himanen, S., Nissinen, A., Auriola, S., Poppy, G., Stewart, C., Holopainen, J. and Nerg, A.-M. (2008). "Constitutive and herbivore-inducible glucosinolate concentrations in oilseed rape (*Brassica napus*) leaves are not affected by Bt Cry1Ac insertion but change under elevated atmospheric CO₂ and O₃." Planta **227**(2): 427-437.
- Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. and Stuber, D. (1988). "Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent." Nat Biotech **6**(11): 1321-1325.
- Hogge, L. R. R., D.W.; Underhill, E.W.; Haughn, G.W. (1988). "HPLC Separation of Glucosinolates from Leaves and Seeds of *Arabidopsis thaliana* and Their Identification Using Thermospray Liquid Chromatography/Mass Spectrometry." Journal of Chromatographic Science **26**: 551-556.
- Hopkins, R. J., van Dam, N. M. and van Loon, J. J. (2009). "Role of glucosinolates in insect-plant relationships and multitrophic interactions." Annu Rev Entomol **54**: 57-83.
- Husebye, H., Chadchawan, S., Winge, P., Thangstad, O. P. and Bones, A. M. (2002). "Guard cell- and phloem idioblast-specific expression of thioglucoside glucohydrolase 1 (myrosinase) in *Arabidopsis*." Plant Physiol **128**(4): 1180-1188.
- Jaillon, O., Aury, J. M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyere, C., Billault, A., Segurens, B., Gouyvenoux, M., Ugarte, E., Cattonaro, F., Anthouard, V., Vico, V., Del Fabbro, C., Alaux, M., Di Gaspero, G., Dumas, V., Felice, N., Paillard, S., Juman, I., Moroldo, M., Scalabrin, S., Canaguier, A., Le Clainche, I., Malacrida, G., Durand, E., Pesole, G., Laucou, V., Chatelet, P., Merdinoglu, D., Delledonne, M., Pezzotti, M., Lecharny, A., Scarpelli, C., Artiguenave, F., Pe, M. E., Valle, G., Morgante, M., Caboche, M., Adam-

- Blondon, A. F., Weissenbach, J., Quetier, F. and Wincker, P. (2007). "The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla." Nature **449**(7161): 463-467.
- Janowitz, T., Trompetter, I. and Piotrowski, M. (2009). "Evolution of nitrilases in glucosinolate-containing plants." Phytochemistry **70**(15-16): 1680-1686.
- Jenrich, R., Trompetter, I., Bak, S., Olsen, C. E., Moller, B. L. and Piotrowski, M. (2007). "Evolution of heteromeric nitrilase complexes in Poaceae with new functions in nitrile metabolism." Proc Natl Acad Sci U S A **104**(47): 18848-18853.
- Jeyaprakash, A. A., Geetha Rani, P., Banuprakash Reddy, G., Banumathi, S., Betzel, C., Sekar, K., Surolia, A. and Vijayan, M. (2002). "Crystal structure of the jacalin-T-antigen complex and a comparative study of lectin-T-antigen complexes." J Mol Biol **321**(4): 637-645.
- Kelly, P. J., Bones, A. and Rossiter, J. T. (1998). "Sub-cellular immunolocalization of the glucosinolate sinigrin in seedlings of *Brassica juncea*." Planta **206**(3): 370-377.
- Kessler, A. and Halitschke, R. (2007). "Specificity and complexity: the impact of herbivore-induced plant responses on arthropod community structure." Current Opinion in Plant Biology **10**(4): 409-414.
- Kimura, M. (1980). "A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences." Journal of Molecular Evolution **16**: 111-120.
- Kissen, R. and Bones, A. M. (2009). "Nitrile-specifier proteins involved in glucosinolate hydrolysis in *Arabidopsis thaliana*." J Biol Chem **284**(18): 12057-12070.
- Kjaer, A. (1973). "The natural distribution of glucosinolates: a uniform group of sulfur containing glucosides." In G. Bendz and J. Santesson [eds.], Chemistry in botanical classification: 229-234.
- Kjær, A. and Olsen Larsen, P. (1976). "Non-protein amino-acids, cyanogenic glycosides, and glucosinolates." Biosynthesis 4, Specialist Periodical Report, Chemical Society, London: 179-203.
- Kliebenstein, D. J., Lambrix, V. M., Reichelt, M., Gershenzon, J. and Mitchell-Olds, T. (2001). "Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in *Arabidopsis*." Plant Cell **13**(3): 681-693.
- Knoop V., M. K. (1980). Gene und Stammbäumer: Ein Handbuch zur molekularen Phylogenetik, Spektrum Akademischer Verlag.
- Koroleva, O. A., Davies, A., Deeken, R., Thorpe, M. R., Tomos, A. D. and Hedrich, R. (2000). "Identification of a new glucosinolate-rich cell type in *Arabidopsis* flower stalk." Plant Physiol **124**(2): 599-608.
- Kroymann, J. (2011). "Natural diversity and adaptation in plant secondary metabolism." Curr Opin Plant Biol **14**(3): 246-251.
- Kuchernig, J. C., Backenkohler, A., Lubbecke, M., Burow, M. and Wittstock, U. (2011). "A thiocyanate-forming protein generates multiple products upon allylglucosinolate breakdown in *Thlaspi arvense*." Phytochemistry **72**(14-15): 1699-1709.

- Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-685.
- Lagercrantz, U. and Lydiate, D. J. (1996). "Comparative genome mapping in Brassica." Genetics **144**(4): 1903-1910.
- Lambrix, V., Reichelt, M., Mitchell-Olds, T., Kliebenstein, D. J. and Gershenzon, J. (2001). "The Arabidopsis epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences Trichoplusia ni herbivory." Plant Cell **13**(12): 2793-2807.
- Leah, R., Kigel, J., Svendsen, I. and Mundy, J. (1995). "Biochemical and Molecular Characterization of a Barley Seed -Glucosidase." Journal of Biological Chemistry **270**(26): 15789-15797.
- Lenman, M., Rodin, J., Josefsson, L. G. and Rask, L. (1990). "Immunological characterization of rapeseed myrosinase." Eur J Biochem **194**(3): 747-753.
- Li, J., Hansen, B. G., Ober, J. A., Kliebenstein, D. J. and Halkier, B. A. (2008). "Subclade of Flavin-Monooxygenases Involved in Aliphatic Glucosinolate Biosynthesis." Plant Physiology **148**(3): 1721-1733.
- Li, Q., Eigenbrode, S. D., Stringam, G. R. and Thiagarajah, M. R. (2000). "Feeding and Growth of Plutella xylostella and Spodoptera eridania on Brassica juncea with Varying Glucosinolate Concentrations and Myrosinase Activities." Journal of Chemical Ecology **26**(10): 2401-2419.
- Lichtenstein, E. P., Strong, F. M. and Morgan, D. G. (1962). "Naturally Occurring Insecticides, Identification of 2-Phenylethylisothiocyanate As an Insecticide Occurring Naturally in the Edible Part of Turnips." Journal of Agricultural and Food Chemistry **10**(1): 30-33.
- Liddle, S., Keresztessy, Z., Hughes, J. and Hughes, M. A. (1998). "A Genomic Cyanogenic Beta-Glucosidase Gene from Cassava (Accession No. X94986). The Electronic Plant Gene Register." Plant Physiology **117**(4): 1525-1528.
- Lüthy, J. and Benn, M. H. (1977). "Thiocyanate formation from glucosinolates: a study of the autolysis of allylglucosinolate in Thlaspi arvense L. seed flour extracts." Canadian Journal of Biochemistry **55**(10): 1028-1031.
- Mandakova, T., Heenan, P. B. and Lysak, M. A. (2010a). "Island species radiation and karyotypic stasis in Pachycladon allopolyploids." BMC Evol Biol **10**: 367.
- Mandakova, T., Joly, S., Krzywinski, M., Mummenhoff, K. and Lysak, M. A. (2010b). "Fast diploidization in close mesopolyploid relatives of Arabidopsis." Plant Cell **22**(7): 2277-2290.
- Mandel, M. and Higa, A. (1970). "Calcium-dependent bacteriophage DNA infection." J Mol Biol **53**(1): 159-162.
- Matile, P. (1980). "The mustard oil bomb - Compartmentation of the myrosinase system." Biochem. Physiol. Pflanzen. **175**: 722-731.
- Matusheski, N. V., Swarup, R., Juvik, J. A., Mithen, R., Bennett, M. and Jeffery, E. H. (2006). "Epithiospecifier protein from broccoli (Brassica oleracea L. ssp. italica) inhibits formation of the anticancer agent sulforaphane." J Agric Food Chem **54**(6): 2069-2076.

- McGadey, J. (1970). "A tetrazolium method for non-specific alkaline phosphatase." Histochemie **23**(2): 180-184.
- Mikkelsen, M. D., Petersen, B. L., Olsen, C. E. and Halkier, B. A. (2002). "Biosynthesis and metabolic engineering of glucosinolates." Amino Acids **22**(3): 279-295.
- Miles, C. I., del Campo, M. L. and Renwick, J. A. (2005). "Behavioral and chemosensory responses to a host recognition cue by larvae of *Pieris rapae*." J Comp Physiol A Neuroethol Sens Neural Behav Physiol **191**(2): 147-155.
- Müller, R., de Vos, M., Sun, J. Y., Sonderby, I. E., Halkier, B. A., Wittstock, U. and Jander, G. (2010). "Differential effects of indole and aliphatic glucosinolates on lepidopteran herbivores." J Chem Ecol **36**(8): 905-913.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction." Cold Spring Harb Symp Quant Biol **51 Pt 1**: 263-273.
- Mumm, R., Burow, M., Bukovinskine'Kiss, G., Kazantzidou, E., Wittstock, U., Dicke, M. and Gershenzon, J. (2008). "Formation of simple nitriles upon glucosinolate hydrolysis affects direct and indirect defense against the specialist herbivore, *Pieris rapae*." Journal of Chemical Ecology **34**(10): 1311-1321.
- Nagano, A. J., Fukao, Y., Fujiwara, M., Nishimura, M. and Hara-Nishimura, I. (2008). "Antagonistic jacalin-related lectins regulate the size of ER body-type beta-glucosidase complexes in *Arabidopsis thaliana*." Plant Cell Physiol **49**(6): 969-980.
- Nielsen, J. K., Dalgaard, L., Larsen, L. M. and SØRensen, H. (1979). "Host plant selection of the horse-radish flea beetle *Phyllotreta armoraciae* (Coleoptera: Chrysomelidae): Feeding responses to glucosinolates from several crucifers." Entomologia Experimentalis et Applicata **25**(3): 227-239.
- Nour-Eldin, H. H., Hansen, B. G., Norholm, M. H., Jensen, J. K. and Halkier, B. A. (2006). "Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments." Nucleic Acids Res **34**(18): e122.
- Ober, D., Harms, R., Witte, L. and Hartmann, T. (2003). "Molecular evolution by change of function. Alkaloid-specific homospermidine synthase retained all properties of deoxyhypusine synthase except binding the eIF5A precursor protein." J Biol Chem **278**(15): 12805-12812.
- Ober, D. and Hartmann, T. (1999a). "Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase." Proc Natl Acad Sci U S A **96**(26): 14777-14782.
- Ober, D. and Hartmann, T. (1999b). "Deoxyhypusine synthase from tobacco. cDNA isolation, characterization, and bacterial expression of an enzyme with extended substrate specificity." J Biol Chem **274**(45): 32040-32047.
- Ober, D. and Hartmann, T. (2000). "Phylogenetic origin of a secondary pathway: the case of pyrrolizidine alkaloids." Plant Mol Biol **44**(4): 445-450.
- Pfalz, M., Vogel, H. and Kroymann, J. (2009). "The gene controlling the indole glucosinolate modifier1 quantitative trait locus alters indole glucosinolate structures and aphid resistance in *Arabidopsis*." Plant Cell **21**(3): 985-999.

- Rask, L., Andreasson, E., Ekbom, B., Eriksson, S., Pontoppidan, B. and Meijer, J. (2000). "Myrosinase: gene family evolution and herbivore defense in Brassicaceae." Plant Mol Biol **42**(1): 93-113.
- Reichelt, M., Brown, P. D., Schneider, B., Oldham, N. J., Stauber, E., Tokuhisa, J., Kliebenstein, D. J., Mitchell-Olds, T. and Gershenzon, J. (2002). "Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*." Phytochemistry **59**(6): 663-671.
- Renwick, J. A. A., Radke, C. D., Sachdev-Gupta, K. and Städler, E. (1992). "Leaf surface chemicals stimulating oviposition by *Pieris rapae* (Lepidoptera: Pieridae) on cabbage." Chemoecology **3**(1): 33-38.
- Rodman, J., Soltis, P., Soltis, D., Sytsma, K. and Karol, K. (1998). "Parallel evolution of glucosinolate biosynthesis inferred from congruent nuclear and plastid gene phylogenies." Am J Bot **85**(7): 997.
- Rodman, J. E. and Chew, F. S. (1980). "Phytochemical correlates of herbivory in a community of native and naturalized cruciferae." Biochemical Systematics and Ecology **8**(1): 43-50.
- Roth, M. J., Tanese, N. and Goff, S. P. (1985). "Purification and characterization of murine retroviral reverse transcriptase expressed in *Escherichia coli*." J Biol Chem **260**(16): 9326-9335.
- Saghai-Maroo, M. A., Soliman, K. M., Jorgensen, R. A. and Allard, R. W. (1984). "Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics." Proc Natl Acad Sci U S A **81**(24): 8014-8018.
- Saitou, N. and Nei, M. (1987). "The neighbor-joining method: a new method for reconstructing phylogenetic trees." Mol Biol Evol **4**(4): 406-425.
- Sarosh, B. R., Wittstock, U., Halkier, B. A. and Ekbom, B. (2010). "The influence of metabolically engineered glucosinolates profiles in *Arabidopsis thaliana* on *Plutella xylostella* preference and performance." Chemoecology **20**(1): 1-9.
- Sharma, N., Cram, D., Huebert, T., Zhou, N. and Parkin, I. A. (2007). "Exploiting the wild crucifer *Thlaspi arvense* to identify conserved and novel genes expressed during a plant's response to cold stress." Plant Mol Biol **63**(2): 171-184.
- Skerra, A. and Schmidt, T. G. (2000). "Use of the Strep-Tag and streptavidin for detection and purification of recombinant proteins." Methods Enzymol **326**: 271-304.
- Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985). "Measurement of protein using bicinchoninic acid." Anal Biochem **150**(1): 76-85.
- Sonderby, I. E., Geu-Flores, F. and Halkier, B. A. (2010). "Biosynthesis of glucosinolates--gene discovery and beyond." Trends Plant Sci **15**(5): 283-290.
- Southern, E. M. (1975). "Detection of specific sequences among DNA fragments separated by gel electrophoresis." J Mol Biol **98**(3): 503-517.
- Spencer, G. F. and Daxenbichler, M. E. (1980). "Gas chromatography-mass spectrometry of nitriles, isothiocyanates and oxazolidinethiones derived from cruciferous glucosinolates." Journal of the Science of Food and Agriculture **31**(4): 359-367.

- Stein, R. J. and Waters, B. M. (2011). "Use of natural variation reveals core genes in the transcriptome of iron-deficient *Arabidopsis thaliana* roots." Journal of Experimental Botany.
- Taipalensuu, J., Eriksson, S. and Rask, L. (1997). "The myrosinase-binding protein from *Brassica napus* seeds possesses lectin activity and has a highly similar vegetatively expressed wound-inducible counterpart." Eur J Biochem **250**(3): 680-688.
- Taji, T., Sakurai, T., Mochida, K., Ishiwata, A., Kurotani, A., Totoki, Y., Toyoda, A., Sakaki, Y., Seki, M., Ono, H., Sakata, Y., Tanaka, S. and Shinozaki, K. (2008). "Large-scale collection and annotation of full-length enriched cDNAs from a model halophyte, *Thellungiella halophila*." BMC Plant Biol **8**: 115.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). "MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods." Mol Biol Evol.
- Textor, S. and Gershenzon, J. (2009). "Herbivore induction of the glucosinolate–myrosinase defense system: major trends, biochemical bases and ecological significance." Phytochemistry Reviews **8**(1): 149-170.
- Thangstad, O., Gilde, B., Chadchawan, S., Seem, M., Husebye, H., Bradley, D. and Bones, A. (2004). "Cell Specific, Cross-Species Expression of Myrosinases in *Brassica Napus*, *Arabidopsis Thaliana* and *Nicotiana Tabacum*." Plant Molecular Biology **54**(4): 597-611.
- Theologis, A., Ecker, J. R., Palm, C. J., Federspiel, N. A., Kaul, S., White, O., Alonso, J., Altafi, H., Araujo, R., Bowman, C. L., Brooks, S. Y., Buehler, E., Chan, A., Chao, Q., Chen, H., Cheuk, R. F., Chin, C. W., Chung, M. K., Conn, L., Conway, A. B., Conway, A. R., Creasy, T. H., Dewar, K., Dunn, P., Etgu, P., Feldblyum, T. V., Feng, J., Fong, B., Fujii, C. Y., Gill, J. E., Goldsmith, A. D., Haas, B., Hansen, N. F., Hughes, B., Huizar, L., Hunter, J. L., Jenkins, J., Johnson-Hopson, C., Khan, S., Khaykin, E., Kim, C. J., Koo, H. L., Kremenetskaia, I., Kurtz, D. B., Kwan, A., Lam, B., Langin-Hooper, S., Lee, A., Lee, J. M., Lenz, C. A., Li, J. H., Li, Y., Lin, X., Liu, S. X., Liu, Z. A., Luros, J. S., Maiti, R., Marziani, A., Militscher, J., Miranda, M., Nguyen, M., Nierman, W. C., Osborne, B. I., Pai, G., Peterson, J., Pham, P. K., Rizzo, M., Rooney, T., Rowley, D., Sakano, H., Salzberg, S. L., Schwartz, J. R., Shinn, P., Southwick, A. M., Sun, H., Tallon, L. J., Tambunga, G., Toriumi, M. J., Town, C. D., Utterback, T., Van Aken, S., Vaysberg, M., Vysotskaia, V. S., Walker, M., Wu, D., Yu, G., Fraser, C. M., Venter, J. C. and Davis, R. W. (2000). "Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana*." Nature **408**(6814): 816-820.
- Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**(22): 4673-4680.
- Tookey, H. L. (1973). "Crambe thioglucoside glucohydrolase (EC 3.2.3.1): separation of a protein required for epithiobutane formation." Can J Biochem **51**(12): 1654-1660.
- Tookey, H. L. and Wolff, I. A. (1970). "Effect of organic reducing agents and ferrous ion on thioglucosidase activity of *Crambe abyssinica* seed." Canadian Journal of Biochemistry **48**(9): 1024-1028.
- Wang, M., Sun, X., Tan, D., Gong, S., Meijer, J. and Zhang, J. (2009). "The two non-functional myrosinase genes TGG3 and TGG6 in *Arabidopsis* are expressed predominantly in pollen." Plant Science **177**(4): 371-375.

- Wentzell, A. M. and Kliebenstein, D. J. (2008). "Genotype, age, tissue, and environment regulate the structural outcome of glucosinolate activation." Plant Physiol **147**(1): 415-428.
- Wheat, C. W., Vogel, H., Wittstock, U., Braby, M. F., Underwood, D. and Mitchell-Olds, T. (2007). "The genetic basis of a plant-insect coevolutionary key innovation." Proc Natl Acad Sci U S A **104**(51): 20427-20431.
- Whelan, S. and Goldman, N. (2001). "A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach." Mol Biol Evol **18**(5): 691-699.
- Winde, I. and Wittstock, U. (2011). "Insect herbivore counteradaptations to the plant glucosinolate-myrosinase system." Phytochemistry **72**(13): 1566-1575.
- Wittstock, U., Agerbirk, N., Stauber, E. J., Olsen, C. E., Hippler, M., Mitchell-Olds, T., Gershenzon, J. and Vogel, H. (2004). "Successful herbivore attack due to metabolic diversion of a plant chemical defense." Proceedings of the National Academy of Sciences of the United States of America **101**(14): 4859-4864.
- Wittstock, U. and Burow, M. (2007). "Tipping the scales--specifier proteins in glucosinolate hydrolysis." IUBMB Life **59**(12): 744-751.
- Wittstock, U., Kliebenstein, D. J., Lambrix, V., Reichelt, M. and Gershenzon, J. (2003). Chapter five Glucosinolate hydrolysis and its impact on generalist and specialist insect herbivores. Recent Advances in Phytochemistry. T. R. John, Elsevier. **Volume 37**: 101-125.
- Xu, Z., Escamilla-Trevino, L., Zeng, L., Lalgondar, M., Bevan, D., Winkel, B., Mohamed, A., Cheng, C. L., Shih, M. C., Poulton, J. and Esen, A. (2004). "Functional genomic analysis of Arabidopsis thaliana glycoside hydrolase family 1." Plant Mol Biol **55**(3): 343-367.
- Xue, J., Jørgensen, M., Pihlgren, U. and Rask, L. (1995). "The myrosinase gene family in Arabidopsis thaliana: gene organization, expression and evolution." Plant Molecular Biology **27**(5): 911-922.
- Yang, Z., Nielsen, R., Goldman, N. and Pedersen, A. M. (2000). "Codon-substitution models for heterogeneous selection pressure at amino acid sites." Genetics **155**(1): 431-449.
- Zdobnov, E. M. and Apweiler, R. (2001). "InterProScan – an integration platform for the signature-recognition methods in InterPro." Bioinformatics **17**(9): 847-848.
- Zhang, J., Pontoppidan, B., Xue, J., Rask, L. and Meijer, J. (2002). "The third myrosinase gene TGG3 in Arabidopsis thaliana is a pseudogene specifically expressed in stamen and petal." Physiol Plant **115**(1): 25-34.
- Zuckerlandl, E. and Pauling, L. (1965). "Molecules as documents of evolutionary history." J Theor Biol **8**(2): 357-366.

7. ANHANG

7.1 Alignments zum Design von degenerierten Primern

<i>A. thaliana</i>	NSP5	1	-----
<i>A. thaliana</i>	At3g07720	1	-----
<i>A. thaliana</i>	NSP1	1	-----
<i>A. thaliana</i>	NSP4	1	MAQKVEAQGGNGGNQWDDGSEYDAVTKIQVAAGGNGIEYVKFTYVKNQQTTEEAPLRGVKG
<i>A. thaliana</i>	NSP3	1	-----
<i>A. thaliana</i>	NSP2	1	-----
<i>B. oleracea</i>	ESP	1	-----
<i>A. thaliana</i>	ESP	1	-----
<i>L. sativum</i>	TFP	1	-----
<i>A. thaliana</i>	NSP5	1	-----
<i>A. thaliana</i>	At3g07720	1	-----
<i>A. thaliana</i>	NSP1	1	-----
<i>A. thaliana</i>	NSP4	61	RSFEADPFVINHPPEEHLVSVEGRYNPEGLILGLTFKSNKTSDLIGYEDGTPFTLQVQDK
<i>A. thaliana</i>	NSP3	1	-----
<i>A. thaliana</i>	NSP2	1	-----
<i>B. oleracea</i>	ESP	1	-----
<i>A. thaliana</i>	ESP	1	-----
<i>L. sativum</i>	TFP	1	-----
<i>A. thaliana</i>	NSP5	1	-----
<i>A. thaliana</i>	At3g07720	1	-----
<i>A. thaliana</i>	NSP1	1	-----MAQKLEAKGGEMGDVWDDGVYENVRKVYVVGQ
<i>A. thaliana</i>	NSP4	121	KIVGFYGFAGNNLHSLGAYFAPLTNVTPLNAKKLEAKGGDTGDIWDDGYVDNVRKVYVVGQ
<i>A. thaliana</i>	NSP3	1	-----MAQKLVAQGGGETGDVWDDGVYDNVTKVYVVGQ
<i>A. thaliana</i>	NSP2	1	-----MVQKVEARGGEIGDVWDDGAYDGVRKVYVVGQ
<i>B. oleracea</i>	ESP	1	-----
<i>A. thaliana</i>	ESP	1	-----
<i>L. sativum</i>	TFP	1	-----
<i>A. thaliana</i>	NSP5	1	-----
<i>A. thaliana</i>	At3g07720	1	-----
<i>A. thaliana</i>	NSP1	32	AQYGIAFVKFEYVNGSQVVVGDEHGKKTELGVVEEFEIDADDYIVYVEGYREKVNDMTS-E
<i>A. thaliana</i>	NSP4	181	AQYGIAFVKFEYVNGSQVVVGDEHGKKTELGVVEEFEIDADDYIVYVEGYREKVNGMTS-E
<i>A. thaliana</i>	NSP3	32	GQYGIAFVKFEYANGSEVVVGDEHGKKTELGVVEEFEIDSDDYIVYVEGYREKVSMTS-E
<i>A. thaliana</i>	NSP2	32	GEDGIAFVKFEYVNGSQEVVGDERGKKTLGAEFEFVDPDDYIVYVEGYHEKVFVTTKE
<i>B. oleracea</i>	ESP	1	-----
<i>A. thaliana</i>	ESP	1	-----
<i>L. sativum</i>	TFP	1	-----
<i>A. thaliana</i>	NSP5	1	-----MCP-VE
<i>A. thaliana</i>	At3g07720	1	-----MAATP-VE
<i>A. thaliana</i>	NSP1	91	MITFLSIKTFKGTSHPIEKRPVKFVLHGGKIVGFHGRSTDVLHSLGAYVSLSSTIKLH
<i>A. thaliana</i>	NSP4	240	MITFLSFKTYKGTSSQPIEQRPQIKFVLQGGKIVGFHGRSTDVLDSLGLAYISLSPTEPLH
<i>A. thaliana</i>	NSP3	91	MITFLSFKTSKGTSSQPIVKKPGVKFVLHGGKIVGFHGRSTDVLHSLGAYVSLSPSTPKLH
<i>A. thaliana</i>	NSP2	92	IISTLTFKTYKGTSPPPGIVSGTKFVLQGGKIVGFHGRSTDVLHSLGAYISSPATPKLR
<i>B. oleracea</i>	ESP	1	-----MAPSVQ
<i>A. thaliana</i>	ESP	1	-----MAPTLQ
<i>L. sativum</i>	TFP	1	-----MALTQ
<i>A. thaliana</i>	NSP5	6	NKWLKVGQKG-AGPGARSSHAMTVVGNKMYCFGGELKPTIHIDNDLYVDFLETQWWSIAP
<i>A. thaliana</i>	At3g07720	8	GKVVQLKQKG-TGPGARSSHAIALVGNKMYAFGGEFQPRVFDNQLYVDFLETQWWSIAP
<i>A. thaliana</i>	NSP1	151	GKWLKVEQKG-EGPGLRCSHGIAQVGNKLYSFGGEFTPNQPIDKHLYVDFLETQWWSISP
<i>A. thaliana</i>	NSP4	300	GKWLKVDENG-DGPGLRCSHGIAQVGNKLYSFGGEFTPNQPIDKHLYVDFLETQWWSISP
<i>A. thaliana</i>	NSP3	151	GKWLKVEQNG-EGPGLRCSHGIAQVGNKLYSFGGELIPNQPIDKHLYVDFLETQWWSIAP
<i>A. thaliana</i>	NSP2	152	GKWLKVEQKG-EGPGPRCSHGIAQVGNKLYSFGGELTPNQPIDKHLYVDFLETQWWSISP
<i>B. oleracea</i>	ESP	7	GEWLKVEQKGGQTGPGPRSSHGIAVVGDKLYSFGGELTPNISIDKDLVDFDNTQWWSISP
<i>A. thaliana</i>	ESP	7	GOWIKVGQKGGTGPGPRSSHGIAVVGDKLYSFGGELTPNKHIDKDLVDFDNTQWWSIAQ
<i>L. sativum</i>	TFP	7	GEWLKVEQKGGPAPGPRSSHCMVVGDKLYMFGGELKPCFHLDKHLYVDFDKNTWWSIAE
<i>A. thaliana</i>	NSP5	65	ATGEAP-FPCFGVSMVTIGSTLYVYGGGRD-DKRRNGLHSYDTEETNEWKLLAPVEE--GL
<i>A. thaliana</i>	At3g07720	67	ASGDAP-PPRVGVMAAVGPIIYFEGGRDSTHOELNELYCFNTLTNQWKLSSGET--GP
<i>A. thaliana</i>	NSP1	210	ATGDVPHHSCLGVRMVSVGSTLYVFGGRD-ASROYNGFYSDFTTNEWKLLTPVEE--GP
<i>A. thaliana</i>	NSP4	359	ATGDIPTTSCLVGMVSVGSTLYVFGGRD-ASROYNGFYSDFTTNEWKLLTPVEE--GP
<i>A. thaliana</i>	NSP3	210	ATGDVPHHSCLGVRMVSVGSTLYVFGGRD-FSROYNGFYSDFTTNEWKLLTPVEE--GP
<i>A. thaliana</i>	NSP2	211	ATGDVPHHSCLGVRMVSVGSSTLYVFGGRD-ASRYNGFYSDFTTNEWKLLTPVEQ--GP
<i>B. oleracea</i>	ESP	67	SKGVAPDVKALGTRMVSVGTRLYLFGGRD-ENKKEDDFYSYDFTTNEWTKLTLLQEGGP
<i>A. thaliana</i>	ESP	67	PKGDAPTISCLGVRMVAVGTRLYLFGGRD-ENRNTEFERSYDFTTSEWFTLTKLLEVGGP
<i>L. sativum</i>	TFP	67	PKGEAPSSISCLGVRMVAVGTRLYLFGGRD-ENRNYSDFYSYDFTVKKEWKLTKLLEERV

Primer 1

Primer 5

<i>A. thaliana</i>	NSP5	121	PCRSYHSMAGD ^{P1} DRKVVYVFGGV ^{P2} -----AKGRVNTLHAYD ^{P5} VVDOKWVEYPAAG ^{P6} ---ACK
<i>A. thaliana</i>	At3g07720	124	QNRSYHS ^{P1} ITADSQNVYVFGGCG ^{P2} -----VDGRLNDLWANYVVDOKWIKFSPG ^{P5} ---ACR
<i>A. thaliana</i>	NSP1	267	TPRS ^{P1} FHSMAD ^{P2} ENVVYVFGGV ^{P5} -----ATARLNTLDSYNIVDKKWFHCSTPG ^{P6} ---SLT
<i>A. thaliana</i>	NSP4	416	TPRS ^{P1} FHSMAD ^{P2} ENVVYVFGGV ^{P5} -----ATARLNTLDSYNIVDKKWFHCSTPG ^{P6} ---SLT
<i>A. thaliana</i>	NSP3	267	TPRS ^{P1} FHSMAD ^{P2} ENVVYVFGGV ^{P5} -----AMDRIKTLDSYNIVDKTWFHCSPNG ^{P6} ---SFS
<i>A. thaliana</i>	NSP2	268	TPRS ^{P1} FHSM ^{P2} AD ^{P5} ENNVYVFGGV ^{P6} -----ATVRLKTLDAYNIVDKWVQCSTPG ^{P6} ---SCS
<i>B. oleracea</i>	ESP	126	EARTYHSMASD ^{P1} ENHVVYVFGGV ^{P2} SKGGTNKTPFRFRTIEAYNIAD ^{P5} GKWSQLPDPG ^{P6} ---QFPRFE
<i>A. thaliana</i>	ESP	126	EARTFHSMA ^{P1} SD ^{P2} ENHVVYVFGGV ^{P5} SKGGTMNTPTRFRTIEAYNIAD ^{P6} GKWAQLPDPG ^{P6} ---NFE
<i>L. sativum</i>	TFP	126	EARSFPA ^{P1} TAAD ^{P2} NHVVY ^{P5} FGGV ^{P6} SKGGVQSTPFRFKSTIVYNIA ^{P6} GTWSQLPNPG ^{P6} ---DFE
<i>A. thaliana</i>	NSP5	172	GRGAPGLV ^{P1} VVEGR ^{P2} IWVLEFGFDGN ^{P5} -----ELGDIHCFDLASEQWKA ^{P6} VETTDVPAAR
<i>A. thaliana</i>	At3g07720	175	GRGGPGL ^{P1} EVVQGL ^{P2} IWVYGFAGE ^{P5} -----EADDVHCEDIAGKEWKEVETK ^{P6} EKPSAR
<i>A. thaliana</i>	NSP1	318	ARGGAGLE ^{P1} VVQGL ^{P2} WVYGFNGC ^{P5} -----EVDDVHYYPD ^{P6} VQDKWTQVET ^{P6} FGVRPSE
<i>A. thaliana</i>	NSP4	467	ARGGAGLE ^{P1} VVQGL ^{P2} WVYGFNGC ^{P5} -----EVDDVHYYPD ^{P6} VQDKWTQVET ^{P6} FGVRPSE
<i>A. thaliana</i>	NSP3	318	IRGGAGLE ^{P1} VVQGL ^{P2} WVYGFNGC ^{P5} -----EVDDVHYYPD ^{P6} PAEDKWTQVET ^{P6} FGVKPNER
<i>A. thaliana</i>	NSP2	319	VRGGAGLE ^{P1} VVQGL ^{P2} WVYGFNGC ^{P5} -----EVDDVHCYDPAQDKWTQVET ^{P6} FGKPCAR
<i>B. oleracea</i>	ESP	186	RRGGAGF ^{P1} VVQGL ^{P2} IWVYGFATSPDPNGKNDY ^{P5} ESDQVQFYDPATQKWTE ^{P6} VETKGDKPSAR
<i>A. thaliana</i>	ESP	183	KRGGAGFA ^{P1} VVQGL ^{P2} IWVYGFATSI ^{P5} VPGGKDDY ^{P6} ESNAVQFYDPASKKWTET ^{P6} ETGAKPSAR
<i>L. sativum</i>	TFP	183	PRGGAGLA ^{P1} VIDK ^{P2} LWVY ^{P5} GFANS-TSGGINDYNSNKVQY ^{P6} YDLVSGKWIE ^{P6} VKTS ^{P6} CVKPSGR
<i>A. thaliana</i>	NSP5	223	SVFPAVS ^{P1} Y ^{P2} GKYI ^{P5} VIYGGEE ^{P6} PELHMHCAC ^{P6} MSGEVY ^{P6} QLDTETL ^{P6} VWE ^{P6} IVCGN-EEEK ^{P6} PS
<i>A. thaliana</i>	At3g07720	226	SVBSTAV ^{P1} VGKQ ^{P2} ILISGGET ^{P5} DPSDLGHMGACCF ^{P6} TGDAY ^{P6} QLDTETL ^{P6} EW ^{P6} WEDGVGSVEHPG
<i>A. thaliana</i>	NSP1	369	SVFASAA ^{P1} GKHI ^{P2} VIFGGET ^{P5} AMDP ^{P6} LAHVGPQ ^{P6} LDGT ^{P6} FALDTETL ^{P6} QWE ^{P6} TDKFGGEEETPS
<i>A. thaliana</i>	NSP4	518	SVFASAA ^{P1} GKHI ^{P2} VIFGGET ^{P5} AMDP ^{P6} LAHVGPQ ^{P6} LDGT ^{P6} FALDTETL ^{P6} QWE ^{P6} TDKFGGEEETPS
<i>A. thaliana</i>	NSP3	369	SVFASAA ^{P1} GKHI ^{P2} VIFGGET ^{P5} AMDP ^{P6} RAHVGPQ ^{P6} LDGT ^{P6} FALDTETL ^{P6} QWE ^{P6} TDKFEG---TPS
<i>A. thaliana</i>	NSP2	370	SVFASAV ^{P1} VGKHI ^{P2} LVFGGEL ^{P5} AMDP ^{P6} KAHEGPQ ^{P6} LSGGT ^{P6} FALDTETL ^{P6} KWE ^{P6} TDKLGEEETPS
<i>B. oleracea</i>	ESP	246	SVFCHAV ^{P1} VGKYL ^{P2} ILIFGGET ^{P5} WPDP ^{P6} KAHLGPGL ^{P6} TSDEGFALDTETL ^{P6} VWE ^{P6} IFGG-GAE---PG
<i>A. thaliana</i>	ESP	243	SVFAHAV ^{P1} VGKYI ^{P2} ILIFAGE ^{P5} WPD ^{P6} LNGHYGPGL ^{P6} TSNEG ^{P6} YALDTETL ^{P6} VWE ^{P6} IFGEEGAP---AI
<i>L. sativum</i>	TFP	242	SVFAYAV ^{P1} VGKQ ^{P2} ILVIYGGEL ^{P5} FRDENGHLGPGL ^{P6} TSNEG ^{P6} YALDTETL ^{P6} VWE ^{P6} IFVDGGE ^{P6} ---MT
<i>A. thaliana</i>	NSP5	282	QRGWCA ^{P1} FTKAV ^{P2} KDGE ^{P5} GLLVHGCNSPTN ^{P6} ERLDD ^{P6} LVFWCF ^{P6} SHLNVN
<i>A. thaliana</i>	At3g07720	286	PRGWCA ^{P1} FAAGSRD ^{P2} GKQGL ^{P5} LVYGCNSP ^{P6} SNDR ^{P6} LDD ^{P6} IEFF ^{P6} TPNTLLK-
<i>A. thaliana</i>	NSP1	429	SRGWTAST ^{P1} TAT ^{P2} TDGKKGL ^{P5} VMHGGKAPT ^{P6} NDRFDD ^{P6} LEFFYGIDSA---
<i>A. thaliana</i>	NSP4	578	SRGWTAST ^{P1} TAT ^{P2} TDGKKGL ^{P5} VMHGGKAPT ^{P6} NDRFDD ^{P6} LEFFYGIDSA---
<i>A. thaliana</i>	NSP3	426	SRGWTAST ^{P1} TCT ^{P2} TDGKKGL ^{P5} VMHGGKAPT ^{P6} NDRFDD ^{P6} LEFFYGIDSV---
<i>A. thaliana</i>	NSP2	430	IRGWSAST ^{P1} TCT ^{P2} TDGKKGL ^{P5} VMHGGKAPT ^{P6} NDRFDD ^{P6} LEFFYGIDSA---
<i>B. oleracea</i>	ESP	302	QLGWPGYT ^{P1} TAT ^{P2} TVYKKGL ^{P5} LMHGGKAPT ^{P6} NDRFDD ^{P6} LEFFYGIDSA---
<i>A. thaliana</i>	ESP	300	PRGWTAT ^{P1} TAAT ^{P2} VDGKNGL ^{P5} LMHGGKAPT ^{P6} NDRFDD ^{P6} LEFFYGIDSA---
<i>L. sativum</i>	TFP	299	PLGWTANCT ^{P1} GTY ^{P2} GKTL ^{P5} GLLMHGGKAPT ^{P6} NDRFDD ^{P6} LEFFYGIDSA---

Abbildung S 1: Positionen der degenerierten Primer. Alignments der Aminosäuresequenzen aller im Vorfeld bekannten spezifizierenden Proteine (Tabelle 1). Erstellt wurde das Multialignment mittels des EBI Webserver für multiple Sequenzalignments ClustalW2. Weiße Buchstaben auf schwarzem Grund zeigen Aminosäuren einer Spalte die in mindestens 80 % der Sequenzen identisch sind. Weiße Buchstaben auf grauem Grund zeigen Aminosäuren einer Spalte die in mindestens 80 % der Sequenzen einer Gruppe mit biochemisch ähnlichen Eigenschaften zugeordnet werden können. Die Positionen der auf Basis der entsprechenden Nukleotidsequenzen generierten P1, P2 (blaue Boxen) P5 und P6 (violette Boxen) sind gekennzeichnet.

7.2 Glucosinolatprofile und -hydrolyseprodukte der untersuchten Spezies

Tabelle S 1: Glucosinolatprofile und Glucosinolat-Hydrolyseprodukte der untersuchten Spezies der Ordnung Brassicales. Glucosinolate wurden mittels HPLC-Analyse als Desulfoglucosinolate bestimmt. Glucosinolat-Hydrolyseprodukte wurden mittels GC-MS-Analyse in Autolysaten analysiert. Die Autolysen wurden von Meike Burow und Loretta Heise sowie im Rahmen der vorliegenden Arbeit durchgeführt. In der dritten Spalte der Tabelle sind unter „Glucosinolate“ die Seitenketten der gefundenen Glucosinolate aufgeführt (4MOI3M = 4-Methoxy-indol-3-yl-methyl, 3MSOP = 3-Methylsulfinylpropyl, 3MTP = 3-Methylthiopropyl, 4MTP = 4-Methylthiopropyl, 4MTB = 4-Methylthiobutyl, I3M = Indol-3-yl-methyl, 4MSOB = 4-Methylsulfinylbutyl, 2-OH-3-butenyl = 2-Hydroxy-3-butenyl, 1MOI3M = 1-Methoxy-indol-3-yl-methyl, 4OHI3M = 4-Hydroxy-indol-3-yl-methyl, 5MSOP = 5-Methylsulfinylpentyl, 9MSON = 9-methylsulfinyl-n-nonyl, 7MSOH = 7-Methylsulfinylheptyl, 8MSOO = 8-Methylsulfinyloctyl, 6MSOH = 6-Methylsulfinylheptyl, 6MTH = 6-Methylthiohexyl, 7MTH = 7-Methylthiohexyl, 8MTO = 8-Methylthiooctyl, 10MSOD = 10-Methylsulfinyldecyl, 3OHP = 3-Hydroxypropyl, 1,4-DiMOI3M = 1,4-Dimethoxy-indol-3-yl-methyl, 11MSON = 11-methylsulfinyl-n-nonyl, 2-OH-2-Phenylethyl = 2-Hydroxy-2-Phenylethyl, pOHB = Para-hydroxy-benzyl)

Spezies	Organ	Glucosinolate	Hydrolyseprodukte
<i>Alliaria petiolata</i> (Sanct Joseph)	Keimlingswurzeln	Allyl, 4MOI3M, 4OHI3M, I3M, Benzyl,	Thiocyanat aus Allylglucosinolat, Epithionitril aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat, Nitril aus Benzylglucosinolat
	Keimblätter	Allyl, 4MOI3M, 4OHI3M, Benzyl	Thiocyanat aus Allylglucosinolat, Epithionitril aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat, Nitril aus Benzylglucosinolat
<i>Alyssum alpestre</i>	Blätter	3MTP, 3MSOP, 4MOI3M, 4MTB, 4MSOB, (unident. Alkyl), I3M, 4MOI3M	Isothiocyanat aus 3-Methylthiopropylglucosinolat, Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat, Nitril aus 3-Methylthiopropylglucosinolat, Nitril aus 3-Methylsulfinylpropylglucosinolat Isothiocyanat aus 3-Methylthiopropylglucosinolat
	Wurzeln	3MTP, 3MSOP, 4MOI3M, 4MTB, I3M 3MTP, 3MSOP	Isothiocyanat aus 3-Methylthiopropylglucosinolat, Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat
	Samen		Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat
<i>Arabis glabra</i>	Blätter	7MSOH, 4MOI3M, 5MSOP	Isothiocyanat aus 7-Methylsulfinylheptylglucosinolat, Isothiocyanat aus 6-Methylsulfinylheptylglucosinolat
	Wurzeln	7MSOH, 7MTH, 6MSOH, 4MOI3M, 6MTH, 4OHI3M, 8MSOO, 5MSOP	Nicht detektierbar
	Samen	7MSOH, 6MSOH, 7MTH	Nicht detektierbar
<i>Barbarea intermedia</i>	Blätter	Benzyl, I3M, 4MOI3M, 2-Phenylethyl	Oxazolidinethion aus 2-Hydroxy-2-Phenylethylglucosinolat Nicht getestet
	Wurzeln	2-Phenylethyl, I3M, (R)-2-OH-2-Phenylethyl, 1MOI3M, 4MOI3M, (S)-2-OH-2-Phenylethyl (R)-2-OH-2-Phenylethyl, (S)-2-OH-Phenylethyl,	Oxazolidinethion aus 2-Hydroxy-2-Phenylethylglucosinolat,

<i>Barbarea stricta</i>	Samen	2-Phenylethyl, I3M	Isothiocyanat aus 2-Phenylethylglucosinolat
	Blätter	(R)-2-OH-2-Phenylethyl, I3M, pOHB	Oxazolidinethion aus 2-Hydroxy-Phenylethylglucosinolat
	Wurzeln	(R)-2-OH-2-Phenylethyl, 2-Phenylethyl I3M, 1MOI3M, (S)-2-OH-Phenylethyl, 4MOI3M	Isothiocyanat aus 2-Phenylethylglucosinolat, Oxazolidinethion aus 2-Hydroxy-Phenylethylglucosinolat Oxazolidinethion aus 2-Hydroxy-Phenylethylglucosinolat
<i>Barbarea vulgaris</i>	Samen	(R)-2-OH-2-Phenylethyl, I3M, pOHB	
	Samen	(R)-2-OH-2-Phenylethyl, 2-Phenylethyl, I3M, (S)-2-OH-Phenylethyl	Oxazolidinethion aus 2-Hydroxy-Phenylethylglucosinolat, Isothiocyanat aus 2-Phenylethylglucosinolat
<i>Camelina microcarpa</i>	Blätter	Nicht detektierbar	Nicht identifizierbar
	Wurzeln	10MSOD, 9MSON, 11MSON, 8MSOO 3MSOP	Nicht identifizierbar
	Blüten	3MSOP, 10MSOD, 9MSON, 3MTP, I3M	Nicht identifizierbar
	Schoten	10MSOD, 9MSON, 11MSON, 8MSOO, 7MSOH	Nicht identifizierbar
<i>Capsella bursa-pastoris</i>	Samen		Nicht identifizierbar
	Blätter	Nicht identifizierbar	Nicht identifizierbar
	Wurzeln	9MSON, 10MSOD, 8MSOO	Nicht identifizierbar
	Blüten	Allyl, 3-Butenyl	Isothiocyanat aus Allylglucosinolat
	Samen	9MSON, 10MSOD, Allyl, 6MSOH, 7MSOH, 8MSOO	Nicht identifizierbar
<i>Capsella rubella</i>	Blätter	10MSOD, 9MSON	Nicht identifizierbar
	Keimlinge	10MSOD, 9MSON	Nicht identifizierbar
	Samen	10MSOD, 9MSON	Nicht getestet
	Blüten		Isothiocyanat aus Allylglucosinolat
<i>Cardamine hirsuta</i>	Blätter	Benzyl, 3-Butenyl, I3M, 4-Pentenyl, 4MSOB, Phenylethyl, 4MOI3M, 4OHI3M, pOHB,	Nitril aus Benzylglucosinolat, Epithionitril aus 3-Butenylglucosinolat (unident. Indolderivat), Isothiocyanat aus Benzylglucosinolat, Nitril aus Phenylethylglucosinolat
	Wurzeln	1MOI3M, Benzyl, I3M, 3-Butenyl, Phenylethyl, I3M, 1,4-DiMOI3M	Epithionitril aus 3-Butenylglucosinolat, Isothiocyanat aus Benzylglucosinolat, Nitril aus Benzylglucosinolat
	Samen	3-Butenyl, Benzyl, 4-Pentenyl, 4OHI3M, I3M	Epithionitril aus 3-Butenylglucosinolat, Isothiocyanat aus Butenylglucosinolat, Isothiocyanat aus Benzylglucosinolat, Isothiocyanat aus 4-Pentenylglucosinolat, Epithionitril aus 4-Pentenylglucosinolat, Nitril aus Benzylglucosinolat
			Epithionitril aus 3-Butenylglucosinolat,
(draußen gesammelt)	Schoten	3-Butenyl, Benzyl, 4-Pentenyl, I3M	

			Isothiocyanat aus 3-Butenylglucosinolat, Isothiocyanat aus Benzylglucosinolat, Isothiocyanat aus 4-Pentenylglucosinolat, Nitril aus Benzylglucosinolat, Epithionitril aus 4-Pentenylglucosinolat
<i>Cardamine impatiens</i>	Blätter	3-Butenyl, I3M, (R)-2-OH-3-Butenyl, (S)-2-OH-3-Butenyl, Benzyl, 4-Pentenyl, (unident. Alkyl)	Epithionitril aus 3-Butenylglucosinolat, Nitril aus Indol-3-yl-methylglucosinolat, Epithionitril aus Butenylglucosinolat
	Samen	3-Butenyl, (S)-2-OH-3-Butenyl, I3M, 4-Pentenyl, (R)-2-OH-3-Butenyl	
<i>Cleome hassleriana</i>	Blätter	Methyl, I3M, Benzyl, 4MOI3M	
	Wurzeln	1,4-Dimethoxy-I3M, I3M, Methyl, 1MOI3M, 4MOI3M	Isothiocyanat aus Butenylglucosinolat, Epithionitril aus Butenylglucosinolat Isothiocyanat aus Methylglucosinolat Isothiocyanat aus Methylglucosinolat
	Samen	Methyl, 2-Methyl-2-OH-Butyl, (R)-2-OH-3-Butenyl, Benzyl	Isothiocyanat aus Methylglucosinolat, Oxazolidinethion aus 2-Methyl-2-Hydroxy-Butylglucosinolat
<i>Cleome spinosa</i>	Blätter	3OHP, 4OHI3M, I3M	Nicht identifizierbar
	Samen	3OHP, Allyl, 4OHI3M, I3M	Nicht getestet
	Blüten	3OHP, 4OHI3M, , I3M, 1MOI3M	Isothiocyanat aus Methylglucosinolat, Isothiocyanat aus Methylpropylglucosinolat, Isothiocyanat aus Methylbutylglucosinolat
<i>Diploaxis muralis</i>	Blätter	bis-4,4-Butyldisulfid, 3-Butenyl, 4MSOB, 4MTB, I3M, 4OHI3M, pOHB	
	Samen	4MTB, 4MSOB	Isothiocyanat aus 4-Methylthiobutylglucosinolat, Isothiocyanat aus 4-Methylsulfinylbutylglucosinolat, Nitril aus 4-Methylthiobutylglucosinolat
	Blüten	4MSOB, 4MTB, bis-4,4-Butyldisulfid, 4OHI3M, I3M, 3-Butenyl	Isothiocyanat aus 4-Methylthiobutylglucosinolat Isothiocyanat aus 4-Methylsulfinylbutylglucosinolat Isothiocyanat aus 4-Methylthiobutylglucosinolat Isothiocyanat aus 4-Mercaptobutylglucosinolat
	Wurzeln	4MTB, 4MOI3M, 4MSOB, 1MOI3M, pOHB, I3M, 4OHI3M	Isothiocyanat aus 4-Methylthiobutylglucosinolat, Nitril aus 4-Methylthiobutylglucosinolat
<i>Draba aurea</i>	Blätter	3-Butenyl, 4OHI3M, Allyl, 4-Pentenyl, 5MSOP, 9MSON, 4MSOB, 4MOI3M, I3M, Phenylethyl, 7MSOH	Epithionitril aus 3-Butenylglucosinolat
	Wurzeln	3-Butenyl, 4OHI3M, I3M, 9MSON	
	Samen	Nicht getestet	Nicht detektierbar Isothiocyanat aus 3-Butenylglucosinolat, Epithionitril aus 3-Butenylglucosinolat
<i>Draba cinerea</i>	Blätter	3MTP, 3MSOP, 4MTB, Methylethyl, (unident. Sulfonyl), 4MSOB, 9MSON, 8MSOO, 6MSOH, 6MTH oder 10MSOD?,	Isothiocyanat aus 3-Methylthiopropylglucosinolat, Isothiocyanat aus 4-Methylthiopropylglucosinolat, Nitril aus 3-Methylthiopropylglucosinolat,

		5MSOP, 7MSOH, 7MTH, 8MTO 4MTB, 3MTP, 9MSON, 4MSOB, 3MSOP	Nitril aus 4-Methylthiopropylglucosinolat,
	Wurzeln		Isothiocyant aus Methylethylglucosinolat Isothiocyant aus 4-Methylthiobutylglucosinolat, Isothiocyant aus 3-Methylthiopropylglucosinolat
<i>Draba lanceolata</i>	Blätter	(R)-2-OH-3-Butenyl, 3-Butenyl, 4MSOB, 5MSOP, 7MSOH, 9MSON, 8MSOO. Allyl, (S)-2-OH-3-Butenyl	Epithionitril aus 2-Hydroxy-3-Butenylglucosinolat, Epithionitril aus 3-Butenylglucosinolat
	Wurzeln	(R)-2-OH-3-Butenyl, 4MTB, 3-Butenyl, 9MSON	Nicht identifizierbar
	Samen	Nicht getestet	Epithionitril aus 3-Butenylglucosinolat, Epithionitril aus 2-Hydroxy-3-Butenylglucosinolat, Oxazolidinethion aus 2-Hydroxy-3-Butenylglucosinolat, Nitril aus 2-Hydroxy-3-Butenylglucosinolat, Isothiocyant aus 3-Butenylglucosinolat
<i>Draba muralis</i>	Blätter	4-Methoxybenzyl	Isothiocyant aus 4-Methoxybenzylglucosinolat, Nitril aus 4-Methoxybenzylglucosinolat
	Wurzeln	4-Methoxybenzyl, pOHb, 4OHl3M	Nicht identifizierbar
	Samen	4-Methoxybenzyl, pOHb, 4OHl3M	Isothiocyant aus 4-Methoxybenzylglucosinolat
<i>Eruca sativa</i>	Blätter	Bis-4,4-Butylgl-disulfid, 3-Butenyl, 4MSOB, 4MOl3M	Isothiocyant aus 2-Mercaptobutylglucosinolat, Isothiocyant aus 4-Methylsulfinylbutylglucosinolat, Isothiocyant aus 4-Methylthiobutylglucosinolat
	Keimlinge	Bis-4,4-Butylgl-disulfid, 4MTB, 4MSOB 4MTB	Nicht getestet
	Samen		Isothiocyant aus 4-Methylthiobutylglucosinolat, Isothiocyant aus 4-Methylsulfinylbutylglucosinolat
	Blüten	4MSOB	Isothiocyant aus 4-Methylsulfinylbutylglucosinolat, Isothiocyant aus 2-Mercaptobutylglucosinolat, Isothiocyant aus 4-Methylthiobutylglucosinolat
	Schoten	4MSOB, Bis-4,4-Butylgl-disulfid, 4MTB 4MSOB, Bis-4,4-Butylgl-disulfid, 4MTB Nicht getestet	Nicht getestet
	Knospen		Nicht getestet
	Wurzeln		Isothiocyant aus 4-Methylthiobutylglucosinolat
<i>Erysimum cheiri</i>	Blätter	3MSOP, 3MTP, 4MOl3M	Isothiocyant aus 3-Methylsulfinylpropylglucosinolat, Isothiocyant aus 3-Methylthiopropylglucosinolat, Isothiocyant aus Allylglucosinolat
	Samen	3-Methylsulfonylpropyl, 3MSOP	Isothiocyant aus 3-Methylsulfonylpropylglucosinolat, Isothiocyant aus 3-Methylsulfinylpropylglucosinolat

<i>Erysimum hieraciifolium</i>	Wurzeln	3MTP, 3MSOP, 4MOI3M,	
	Blätter	(unident. Alkyl), Allyl, 3MSOP, 4MOI3M, 4OHI3M, 4MSOB, I3M	Isothiocyanat aus Allylglucosinolat
	Samen	3-Methylsulfonylpropyl, Allyl, 3MSOP, 4OHI3M	Isothiocyanat aus 3-Methylsulfonylpropylglucosinolat, Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat, Isothiocyanat aus Allylglucosinolat
	Wurzeln	3MTP, Allyl, 3MSOP, (unident. Alkyl), 4MOI3M, 4MTB, 4MSOB	Isothiocyanat aus 3-Methylthiopropylglucosinolat, Nitril aus 3-Methylthiopropylglucosinolat
<i>Iberis amara</i>	Blätter	3MSOP, 3MTP, 4MOI3M, 4MSOB	Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat
	Samen	3MSOP, 3MTP	Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat
	Blüten	Nicht getestet	Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat
	Wurzeln	3MSOP, 3MTP, 4MOI3M, 4MSOB	Nitril aus 3-Methylthiopropylglucosinolat, Isothiocyanat aus 3-Methylthiopropylglucosinolat, Isothiocyanat aus 3-Methylsulfinylpropylglucosinolat
<i>Isatis tinctoria</i>	Blätter	Nicht getestet	Nitril aus Indol-3-yl-methylglucosinolat
	Keimlinge	(R)-2-OH-3-Butenyl, 1MOI3M, I3M, (S)-2-OH-3-Butenyl, 3-Butenyl	Nicht getestet
	Samen	(R)-2-OH-3-Butenyl, (S)-2-OH-3-Butenyl, I3M, 1MOI3M, 3-Butenyl	Nicht identifizierbar
	Wurzeln	Nicht getestet	
<i>Limnanthes douglasii</i>	Blätter	I3M, 4-Methoxybenzyl	Nitril aus Indol-3-yl-methylglucosinolat
	Samen	4-Methoxybenzyl	Isothiocyanat aus 4-Methoxybenzylglucosinolat, Isothiocyanat aus Para-hydroxy-benzylglucosinolat
	Wurzeln	4-Methoxybenzyl	Isothiocyanat aus 4-Methoxybenzylglucosinolat
	Blätter	pOHB, Benzyl, Allyl	Isothiocyanat aus 4-Methoxybenzylglucosinolat
<i>Moringa oleifera</i>	Samen	Rhamnopyranosyloxybenzyl, pOHB	Isothiocyanat aus Benzylglucosinolat
	Wurzeln		Nicht getestet
<i>Reseda lutea</i>	Blätter	Benzyl, I3M	Nicht identifizierbar
	Samen	Benzyl, pOHB, I3M	Nicht identifizierbar
	Wurzeln	Benzyl, I3M, Phenylethyl, 4OHI3M, pOHB	Isothiocyanat aus Benzylglucosinolat
	Blätter	(S)-OH-2-Phenylethyl, I3M, (R)-OH-2-Phenylethyl, Phenylethyl, 4MOI3M	Oxazolidinethion aus 2-Hydroxy-2-Phenylethylglucosinolat
<i>Reseda luteola</i>	Samen	Nicht getestet	Oxazolidinethion aus 2-Hydroxy-2-Phenylethylglucosinolat
	Wurzeln	Phenylethyl, I3M, (S)-OH-2-Phenylethyl	Isothiocyanat aus Phenylethylglucosinolat, Oxazolidinethion aus 2-Hydroxy-2-Phenylethylglucosinolat
	Blätter		
	Keimlinge		
<i>Schouwia purpurea</i>	Blätter	Nicht getestet	Epithionitril aus 3-Butenylglucosinolat, Nitril aus Indol-3-yl-methylglucosinolat
	Wurzeln		Epithionitril aus Butenylglucosinolat,
	Keimlinge	(S)-2-OH-3-Butenyl, 1MOI3M, 3-Butenyl,	

<i>Thlaspi arvense</i>		4MOI3M, I3M, 3OHP, (R)-2-OH-3-Butenyl	Epithionitril aus 3-Butenylglucosinolat, Nitril aus Indol-3-yl-methylglucosinolat
	Blätter	Allyl, Benzyl, 3MSOP, 4MOI3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat,
	Keimlinge	Allyl, 1MOI3M, 4MOI3M, Benzyl, I3M, 3MSOP, 4OHI3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat,
	Samen	Allyl, Benzyl, I3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat
	Wurzeln	Allyl, 1MOI3M, 4MOI3M, Benzyl, I3M	Thiocyanat aus Allylglucosinolat, Isothiocyanat aus Benzylglucosinolat Nitril aus Benzylglucosinolat
	Blüten	Allyl, Benzyl, 4MOI3M, 4OHI3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat, Nitril aus Allylglucosinolat
	Schoten	Allyl, 3MSOP, Benzyl, 4OHI3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat, Isothiocyanat aus Allylglucosinolat
<i>Tropaeolum minus</i>	Stängel	Allyl, 4MOI3M, Benzyl, 4OHI3M	Isothiocyanat aus Allylglucosinolat Epithionitril aus Allylglucosinolat, Thiocyanat aus Allylglucosinolat,
	Blätter	Nicht getestet	Isothiocyanat aus Benzylglucosinolat, Nitril aus Benzylglucosinolat
	Samen Blüten	Benzyl, 4-Methoxybenzyl, pOHB Benzyl, pOHB	Nicht getestet Isothiocyanat aus Benzylglucosinolat, Nitril aus Benzylglucosinolat

7.3 Paarweiser Identitätsvergleich aller Sequenzen spezifizierender Proteine mittels ClustalW

Tabelle S 2: Identitätsvergleich aller Sequenzen spezifizierender Proteine. Die in Prozent (%) angegebenen paarweise ermittelten Werte wurden mittels des EBI Webservers für multiple Sequenzalignments ermittelt. Sie sind definiert als der Quotient aus der Anzahl an Identitäten des besten Alignments und der Gesamtanzahl verglichener Positionen (Lückenpositionen werden ausgeschlossen). Oben rechts abgebildet finden sich die Daten der Aminosäuresequenzidentität, unten links diejenigen der Nukleotididentität.

Benennung	<i>Lepidium sativum</i> (Lepidieae) LsTFP	<i>Alliaria petiolata</i> (Thlaspiaceae) TFP	<i>Thlaspi arvense</i> (Thlaspiaceae) TaTFP	<i>Draba aurea</i> (Arabideae) ESP	<i>Draba lanceolata</i> (Arabideae) ESP	<i>Brassica oleracea</i> (Brassicaceae) ESP	<i>Brassica rapa</i> (Brassicaceae) FJ37463	<i>Schouwia purpurea</i> (Brassicaceae) ESP	<i>Arabidopsis lyrata</i> (Camelineae) XM002884120	<i>Arabidopsis thaliana</i> (Camelineae) AtESP	<i>Cardamine hirsuta</i> (Cardamineae) ESP	<i>Cardamine impatiens</i> (Cardamineae) ESP	<i>Isatis tinctoria</i> (Isatideae) ESP	<i>Schouwia purpurea</i> (Brassicaceae) NSP	<i>Arabidopsis thaliana</i> (Camelineae) ANSP1	<i>Arabidopsis thaliana</i> (Camelineae) ANSP2	<i>Arabidopsis thaliana</i> (Camelineae) ANSP3	<i>Arabidopsis thaliana</i> (Camelineae) ANSP4	<i>Arabidopsis thaliana</i> (Camelineae) ANSP5	<i>Arabidopsis thaliana</i> (Camelineae) At3g07720	<i>Cardamine hirsuta</i> (Cardamineae) NSP	<i>Eutrema halophilum</i> (Eutremeae) AK353179	<i>Isatis tinctoria</i> (Isatideae) NSP
<i>Lepidium sativum</i> (Lepidieae) LsTFP		60	59	67	67	63	64	65	67	68	65	64	63	51	51	49	51	48	46	44	49	45	48
<i>Alliaria petiolata</i> (Thlaspiaceae) TFP	73		93	74	73	75	75	76	72	72	68	67	76	53	53	54	52	50	46	47	52	46	52
<i>Thlaspi arvense</i> (Thlaspiaceae) TaTFP	73	94		72	72	72	73	73	70	70	67	65	75	52	52	63	51	50	46	46	51	47	52
<i>Draba aurea</i> (Arabideae) ESP	75	80	79		98	79	76	78	78	79	74	72	81	57	57	57	56	55	47	49	55	48	55
<i>Draba lanceolata</i> (Arabideae) ESP	75	80	80	98		80	76	78	78	78	74	72	81	58	56	57	56	55	47	48	54	48	55
<i>Brassica oleracea</i> (Brassicaceae) ESP	74	80	79	81	81		85	89	76	77	73	71	89	56	55	57	54	53	48	47	53	48	54
<i>Brassica rapa</i> (Brassicaceae) FJ37463	76	82	81	80	80	87		85	77	76	75	72	84	53	55	54	53	53	47	48	53	47	53
<i>Schouwia purpurea</i> (Brassicaceae) ESP	77	82	82	82	82	88	87		76	78	73	71	86	56	57	56	55	54	46	49	54	48	55
<i>Arabidopsis lyrata</i> (Camelineae) XM002884120	78	78	78	80	80	79	81	81		83	78	76	79	57	55	57	54	53	47	46	53	49	55
<i>Arabidopsis thaliana</i> (Camelineae) AtESP	78	79	78	81	81	80	81	81	85		78	76	78	56	56	56	55	54	50	48	53	51	54
<i>Cardamine hirsuta</i> (Cardamineae) ESP	77	77	76	79	79	79	80	80	83	82		85	76	54	52	53	51	50	48	47	51	48	51
<i>Cardamine impatiens</i> (Cardamineae) ESP	74	77	76	78	79	78	79	80	83	83	89		73	54	53	54	51	51	48	47	51	49	52
<i>Isatis tinctoria</i> (Isatideae) ESP	75	83	83	84	84	88	87	88	81	82	80	81		58	57	60	57	55	48	48	55	49	57
<i>Schouwia purpurea</i> (Brassicaceae) NSP	67	65	64	68	69	66	65	68	68	65	67	66	65		80	80	79	77	53	52	81	54	83
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	67	63	63	67	67	66	64	67	66	66	64	64	66	85		81	90	92	53	51	84	53	77
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	64	66	66	68	68	64	66	66	67	66	65	64	65	85	86		79	79	52	52	82	54	75
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	68	63	64	66	67	63	64	65	67	67	64	63	65	83	92	85		85	50	49	81	52	74
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	67	64	64	67	68	68	66	66	67	67	65	65	67	83	95	85	90		52	50	81	52	74
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP5	63	63	63	66	65	64	63	61	64	64	64	61	64	67	67	68	66	67		54	51	85	52
<i>Arabidopsis thaliana</i> (Camelineae) At3g07720	62	62	63	61	61	64	62	65	63	64	65	64	62	67	66	66	66	66	65		51	57	51
<i>Cardamine hirsuta</i> (Cardamineae) NSP	65	63	62	67	67	65	65	65	68	67	65	64	66	84	88	86	86	87	66	66		51	75
<i>Eutrema halophilum</i> (Eutremeae) AK353179	62	62	64	64	65	66	65	63	63	63	65	61	62	68	67	68	67	70	84	65	66		53
<i>Isatis tinctoria</i> (Isatideae) NSP	65	65	64	67	66	66	65	67	67	67	67	65	65	87	83	82	82	82	67	67	83	68	

7.4 Datenmatrizen der Berechnung von d_{aa} und d_{na}

Tabelle S 3: Aminosäure- und Nukleotiddistanz spezifizierender Proteine. Tabellarisch aufgestellt wurden alle paarweise verglichenen Aminosäure- (d_{aa}) sowie Nukleotiddistanzen (d_{na}) definiert als die Anzahl an Aminosäuresubstitutionen / Nukleotidsubstitutionen über die Länge des gesamten Alignments. Berechnungsgrundlage war das Equal-input Modell ergänzt durch positionsspezifisch variable Raten (+G) für d_{aa} sowie das Kimura-2-Parameter-Modell für d_{na} . Oben rechts abgebildet finden sich die Daten der Aminosäuredistanz, unten links diejenigen der Nukleotiddistanz.

Benennung	<i>Lepidium sativum</i> (Lepideae) LsTFP	<i>Alliaria petiolata</i> (Thlaspideae) TFP	<i>Thlaspi arvense</i> (Thlaspideae) TaTFP	<i>Draba aurea</i> (Arabideae) ESP	<i>Draba lanceolata</i> (Arabideae) ESP	<i>Brassica oleracea</i> (Brassicaceae) ESP	<i>Brassica rapa</i> (Brassicaceae) FJ37463	<i>Schouwia purpurea</i> (Brassicaceae) ESP	<i>Arabidopsis lyrata</i> (Camelineae) XM002884120	<i>Arabidopsis thaliana</i> (Camelineae) AtESP	<i>Cardamine hirsuta</i> (Cardamineae) ESP	<i>Cardamine impatiens</i> (Cardamineae) ESP	<i>Isatis tinctoria</i> (Isatideae) ESP	<i>Schouwia purpurea</i> (Brassicaceae) NSP	<i>Arabidopsis thaliana</i> (Camelineae) ANSP1	<i>Arabidopsis thaliana</i> (Camelineae) ANSP2	<i>Arabidopsis thaliana</i> (Camelineae) ANSP3	<i>Arabidopsis thaliana</i> (Camelineae) ANSP4	<i>Arabidopsis thaliana</i> (Camelineae) ANSP5	<i>Arabidopsis thaliana</i> (Camelineae) At3g07720	<i>Cardamine hirsuta</i> (Cardamineae) NSP	<i>Eutrema halophilum</i> (Eutremeae) AK353179	<i>Isatis tinctoria</i> (Isatideae) NSP
<i>Lepidium sativum</i> (Lepideae) LsTFP		0,681	0,701	0,503	0,495	0,605	0,578	0,561	0,497	0,479	0,560	0,605	0,587	1,010	1,024	1,071	0,985	1,151	1,461	1,461	1,102	1,508	1,136
<i>Alliaria petiolata</i> (Thlaspideae) TFP	0,531		0,065	0,335	0,341	0,323	0,317	0,299	0,374	0,373	0,434	0,495	0,282	0,812	0,813	0,790	0,815	0,926	1,336	1,316	0,861	1,399	0,838
<i>Thlaspi arvense</i> (Thlaspideae) TaTFP	0,545	0,059		0,347	0,354	0,366	0,341	0,341	0,408	0,406	0,464	0,527	0,299	0,836	0,848	0,824	0,839	0,952	1,335	1,396	0,873	1,317	0,862
<i>Draba aurea</i> (Arabideae) ESP	0,419	0,290	0,303		0,010	0,250	0,294	0,266	0,272	0,266	0,348	0,393	0,219	0,713	0,693	0,693	0,704	0,756	1,262	1,172	0,757	1,264	0,815
<i>Draba lanceolata</i> (Arabideae) ESP	0,416	0,292	0,305	0,008		0,244	0,294	0,266	0,267	0,266	0,335	0,380	0,219	0,703	0,713	0,693	0,694	0,779	1,243	1,207	0,779	1,263	0,815
<i>Brassica oleracea</i> (Brassicaceae) ESP	0,445	0,274	0,291	0,271	0,270		0,185	0,111	0,312	0,294	0,354	0,414	0,123	0,735	0,724	0,703	0,747	0,802	1,207	1,281	0,826	1,282	0,827
<i>Brassica rapa</i> (Brassicaceae) FJ37463	0,439	0,234	0,252	0,275	0,274	0,169		0,166	0,295	0,299	0,311	0,380	0,180	0,825	0,734	0,790	0,780	0,813	1,242	1,171	0,813	1,359	0,876
<i>Schouwia purpurea</i> (Brassicaceae) ESP	0,400	0,232	0,242	0,234	0,229	0,124	0,155		0,312	0,271	0,354	0,407	0,153	0,745	0,682	0,703	0,725	0,768	1,319	1,138	0,779	1,301	0,791
<i>Arabidopsis lyrata</i> (Camelineae) XM002884120	0,353	0,302	0,332	0,268	0,262	0,295	0,254	0,254		0,200	0,278	0,312	0,245	0,749	0,818	0,717	0,808	0,905	1,232	1,347	0,867	1,215	0,831
<i>Arabidopsis thaliana</i> (Camelineae) AtESP	0,345	0,309	0,333	0,246	0,245	0,279	0,258	0,234	0,184		0,282	0,311	0,266	0,790	0,779	0,756	0,781	0,850	1,073	1,206	0,850	1,090	0,814
<i>Cardamine hirsuta</i> (Cardamineae) ESP	0,396	0,360	0,381	0,299	0,295	0,293	0,271	0,256	0,219	0,248		0,162	0,299	0,824	0,900	0,862	0,890	0,981	1,187	1,278	0,940	1,242	0,941
<i>Cardamine impatiens</i> (Cardamineae) ESP	0,408	0,357	0,389	0,303	0,299	0,319	0,285	0,255	0,216	0,213	0,125		0,354	0,837	0,887	0,850	0,904	0,982	1,188	1,260	0,954	1,190	0,955
<i>Isatis tinctoria</i> (Isatideae) ESP	0,411	0,196	0,222	0,213	0,211	0,132	0,141	0,120	0,230	0,228	0,257	0,248		0,662	0,662	0,605	0,663	0,745	1,206	1,206	0,713	1,244	0,724
<i>Schouwia purpurea</i> (Brassicaceae) NSP	0,875	0,734	0,764	0,665	0,651	0,697	0,730	0,637	0,702	0,749	0,686	0,787	0,676		0,185	0,157	0,201	0,234	0,955	0,983	0,153	0,956	0,153
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	0,853	0,772	0,796	0,653	0,653	0,711	0,696	0,656	0,730	0,761	0,690	0,812	0,659	0,166		0,194	0,100	0,057	0,983	1,012	0,162	0,984	0,229
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	0,957	0,707	0,730	0,666	0,673	0,741	0,699	0,679	0,744	0,744	0,711	0,807	0,690	0,157	0,145		0,232	0,229	1,012	0,969	0,180	0,956	0,239
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	0,888	0,729	0,752	0,662	0,657	0,762	0,733	0,692	0,717	0,772	0,709	0,825	0,664	0,181	0,085	0,160		0,155	1,047	1,095	0,216	1,033	0,286
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	0,889	0,800	0,840	0,667	0,667	0,739	0,730	0,687	0,769	0,788	0,727	0,849	0,683	0,191	0,041	0,155	0,108		1,026	1,104	0,214	1,012	0,271
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP5	1,159	1,189	1,168	0,971	0,994	1,045	1,044	1,061	1,040	1,034	1,054	1,153	1,023	0,886	0,818	0,827	0,893	0,838		0,928	1,073	0,149	0,998
<i>Arabidopsis thaliana</i> (Camelineae) At3g07720	1,211	1,230	1,205	0,971	0,996	1,018	1,046	0,943	0,982	1,090	1,061	1,086	1,083	0,815	0,800	0,783	0,844	0,845	0,872		1,012	0,827	1,028
<i>Cardamine hirsuta</i> (Cardamineae) NSP	0,935	0,757	0,790	0,662	0,665	0,760	0,729	0,703	0,725	0,760	0,751	0,890	0,687	0,177	0,127	0,154	0,155	0,140	0,828	0,820		1,074	0,229
<i>Eutrema halophilum</i> (Eutremeae) AK353179	1,238	1,178	1,096	1,040	1,033	1,066	1,081	1,076	1,039	1,077	1,031	1,126	1,078	0,864	0,864	0,794	0,894	0,920	0,215	0,853	0,852		0,971
<i>Isatis tinctoria</i> (Isatideae) NSP	1,014	0,748	0,767	0,728	0,740	0,733	0,768	0,692	0,744	0,786	0,743	0,848	0,708	0,138	0,188	0,205	0,212	0,211	0,901	0,812	0,193	0,891	

Tabelle S 4: Aminosäure- und Nukleotiddistanz spezifizierender Proteine. Tabellarisch aufgestellt wurden die paarweisen Vergleiche der Aminosäure- (d_{aa}) sowie Nukleotiddistanzen (d_{na}) definiert als die Anzahl an Aminosäuresubstitutionen / Nukleotidsubstitutionen über die Länge des gesamten Alignments. Die Berechnung wurde beschränkt auf die Kelchdomänen derjenigen NSPs mit Lektin-ähnlichen Jacalindomänen. Berechnungsgrundlage war das Equal-input Modell ergänzt durch positionsspezifisch variable Raten (+G) für d_{aa} sowie das Kimura-2-Parameter-Modell für d_{na} . Oben rechts abgebildet finden sich die Daten der Aminosäuredistanz, unten links diejenigen der Nukleotiddistanz.

Benennung	<i>Schouwia purpurea</i> (Brassicaceae) NSP	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	<i>Cardamine hirsuta</i> (Cardamineae) NSP	<i>Isatis tinctoria</i> (Isatideae) NSP
<i>Schouwia purpurea</i> (Brassicaceae) NSP		0,184	0,157	0,203	0,233	0,148	0,148
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	0,163		0,194	0,098	0,057	0,161	0,223
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	0,154	0,145		0,228	0,228	0,179	0,233
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	0,183	0,087	0,162		0,152	0,218	0,281
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	0,188	0,041	0,155	0,109		0,213	0,265
<i>Cardamine hirsuta</i> (Cardamineae) NSP	0,175	0,126	0,152	0,155	0,138		0,223
<i>Isatis tinctoria</i> (Isatideae) NSP	0,138	0,188	0,205	0,213	0,211	0,191	

Tabelle S 5: Aminosäure- und Nukleotiddistanz spezifizierender Proteine. Tabellarisch aufgestellt wurden die paarweisen Vergleiche der Aminosäure- (d_{aa}) sowie Nukleotiddistanzen (d_{na}) definiert als die Anzahl an Aminosäuresubstitutionen / Nukleotidsubstitutionen über die Länge des gesamten Alignments. Die Berechnung wurde beschränkt auf die Lektin-ähnlichen Jacalindomänen der NSPs. Hierbei wurden die zwei Domänen des AtNSP4 getrennt betrachtet. Die N-terminal gelegene Domäne wurde hierfür als AtNSP4A und die folgende als AtNSP4B bezeichnet. Berechnungsgrundlage war das Equal-input Modell ergänzt durch positionsspezifisch variable Raten (+G) für d_{aa} sowie das Kimura-2-Parameter-Modell für d_{na} . Oben rechts abgebildet finden sich die Daten der Aminosäuredistanz, unten links diejenigen der Nukleotiddistanz.

Benennung	<i>Schouwia purpurea</i> (Brassicaceae) NSP	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4A	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4B	<i>Cardamine hirsuta</i> (Cardamineae) NSP	<i>Isatis tinctoria</i> (Isatideae) NSP
<i>Schouwia purpurea</i> (Brassicaceae) NSP		0,409	0,516	0,443	2,318	0,496	0,409	0,361
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	0,267		0,332	0,128	1,974	0,119	0,203	0,533
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	0,311	0,220		0,362	1,640	0,332	0,277	0,710
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	0,287	0,087	0,258		1,905	0,180	0,251	0,639
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4A	1,863	1,758	1,558	1,673		1,972	1,764	2,497
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4B	0,313	0,083	0,244	0,102	1,909		0,203	0,594
<i>Cardamine hirsuta</i> (Cardamineae) NSP	0,279	0,134	0,213	0,155	1,579	0,137		0,595
<i>Isatis tinctoria</i> (Isatideae) NSP	0,207	0,300	0,348	0,343	2,011	0,316	0,308	

7.5 Datenmatrizen der Berechnung von $\omega = d_N / d_S$

Tabelle S 6: Substitutionsratenverhältnisse (ω) spezifizierender Proteine. Tabellarisch dargestellt sind die Substitutionsratenverhältnisse ω definiert als $\omega = d_N / d_S$ (Quotient aus d_N und d_S) mit d_N als Anzahl nichtsynonymer Substitutionen pro nichtsynonymer Sequenzposition und d_S als Anzahl synonyme Substitutionen pro synonyme Sequenzposition berechnet unter Verwendung des Kimura 2-parameter-Modell.

Benennung	<i>Lepidium sativum</i> (Lepideae) LsTFP	<i>Alliaria petiolata</i> (Thlaspeideae) TFP	<i>Thlaspi arvense</i> (Thlaspeideae) TaTFP	<i>Draba aurea</i> (Arabideae) ESP	<i>Draba lanceolata</i> (Arabideae) ESP	<i>Brassica oleracea</i> (Brassicaceae) ESP	<i>Brassica rapa</i> (Brassicaceae) FJ37463	<i>Schouwia purpurea</i> (Brassicaceae) ESP	<i>Arabidopsis lyrata</i> (Camelineae) XM002884120	<i>Arabidopsis thaliana</i> (Camelineae) AtESP	<i>Cardamine hirsuta</i> (Cardamineae) ESP	<i>Cardamine impatiens</i> (Cardamineae) ESP	<i>Isatis tinctoria</i> (Isatideae) ESP	<i>Schouwia purpurea</i> (Brassicaceae) NSP	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP5	<i>Arabidopsis thaliana</i> (Camelineae) At3g07720	<i>Cardamine hirsuta</i> (Cardamineae) NSP	<i>Eutrema halophilum</i> (Eutremeae) AK353179	<i>Isatis tinctoria</i> (Isatideae) NSP
<i>Lepidium sativum</i> (Lepideae) LsTFP																							
<i>Alliaria petiolata</i> (Thlaspeideae) TFP	0,516																						
<i>Thlaspi arvense</i> (Thlaspeideae) TaTFP	0,524	0,198																					
<i>Draba aurea</i> (Arabideae) ESP	0,380	0,348	0,344																				
<i>Draba lanceolata</i> (Arabideae) ESP	0,385	0,352	0,352	0,627																			
<i>Brassica oleracea</i> (Brassicaceae) ESP	0,397	0,359	0,392	0,218	0,214																		
<i>Brassica rapa</i> (Brassicaceae) FJ37463	0,412	0,403	0,381	0,277	0,265	0,213																	
<i>Schouwia purpurea</i> (Brassicaceae) ESP	0,460	0,446	0,478	0,285	0,293	0,198	0,222																
<i>Arabidopsis lyrata</i> (Camelineae) XM002884120	0,427	0,444	0,415	0,279	0,285	0,304	0,339	0,360															
<i>Arabidopsis thaliana</i> (Camelineae) AtESP	0,449	0,383	0,377	0,287	0,297	0,265	0,305	0,285	0,258														
<i>Cardamine hirsuta</i> (Cardamineae) ESP	0,405	0,428	0,417	0,379	0,364	0,321	0,281	0,388	0,322	0,325													
<i>Cardamine impatiens</i> (Cardamineae) ESP	0,495	0,549	0,511	0,478	0,475	0,381	0,416	0,549	0,434	0,478	0,394												
<i>Isatis tinctoria</i> (Isatideae) ESP	0,448	0,502	0,439	0,273	0,269	0,143	0,265	0,287	0,331	0,327	0,341	0,522											
<i>Schouwia purpurea</i> (Brassicaceae) NSP	0,312	0,312	0,333	0,240	0,241	0,250	0,276	0,299	0,264	0,244	0,278	0,253	0,236										
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	0,309	0,306	0,308	0,248	0,254	0,233	0,280	0,262	0,277	0,242	0,294	0,260	0,253	0,300									
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	0,251	0,320	0,342	0,262	0,254	0,225	0,271	0,245	0,236	0,238	0,293	0,251	0,219	0,288	0,356								
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	0,295	0,342	0,346	0,250	0,249	0,236	0,281	0,277	0,294	0,250	0,302	0,285	0,278	0,291	0,308	0,351							
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	0,335	0,321	0,323	0,255	0,261	0,253	0,289	0,282	0,282	0,246	0,307	0,279	0,270	0,316	0,319	0,349	0,418						
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP5	0,312	0,250	0,238	0,328	0,326	0,272	0,252	0,297	0,269	0,222	0,272	0,262	0,292	0,239	0,300	0,289	0,307	0,296					
<i>Arabidopsis thaliana</i> (Camelineae) At3g07720	0,244	0,234	0,288	0,223	0,240	0,316	0,275	0,302	0,351	0,239	0,311	0,269	0,236	0,306	0,303	0,295	0,323	0,308	0,270				
<i>Cardamine hirsuta</i> (Cardamineae) NSP	0,302	0,328	0,335	0,259	0,264	0,256	0,298	0,271	0,303	0,264	0,316	0,270	0,251	0,259	0,343	0,344	0,429	0,435	0,322	0,330			
<i>Eutrema halophilum</i> (Eutremeae) AK353179	0,269	0,287	0,278	0,276	0,285	0,290	0,270	0,296	0,281	0,225	0,315	0,263	0,273	0,294	0,269	0,300	0,305	0,255	0,123	0,256	0,319		
<i>Isatis tinctoria</i> (Isatideae) NSP	0,281	0,306	0,332	0,255	0,246	0,281	0,296	0,292	0,271	0,231	0,312	0,279	0,266	0,360	0,324	0,348	0,383	0,357	0,277	0,298	0,372	0,302	

Tabelle S 7: Substitutionsratenverhältnisse (ω) spezifizierender Proteine Tabellarisch dargestellt sind die Substitutionsratenverhältnisse ω definiert als $\omega = d_N / d_S$ (Quotient aus d_N und d_S) mit d_N als Anzahl nichtsynonymer Substitutionen pro nichtsynonymer Sequenzposition und d_S als Anzahl synonyme Substitutionen pro synonymer Sequenzposition berechnet unter Verwendung des Kimura-2-Parameter-Modell. Die Berechnung wurde beschränkt auf die Kelchdomänen derjenigen NSPs mit Lektin-ähnlichen Jacalindomänen.

Benennung	<i>Schouwia purpurea</i> (Brassicaceae) NSP	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	<i>Cardamine hirsuta</i> (Cardamineae) NSP	<i>Isatis tinctoria</i> (Isatideae) NSP
<i>Schouwia purpurea</i> (Brassicaceae) NSP							
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	0,236						
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	0,202	0,334					
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	0,226	0,310	0,345				
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4	0,270	0,378	0,357	0,410			
<i>Cardamine hirsuta</i> (Cardamineae) NSP	0,200	0,358	0,329	0,445	0,475		
<i>Isatis tinctoria</i> (Isatideae) NSP	0,233	0,259	0,261	0,315	0,293	0,290	

Tabelle S 8: Substitutionsratenverhältnisse (ω) spezifizierender Proteine Tabellarisch dargestellt sind die Substitutionsratenverhältnisse ω definiert als $\omega = d_N / d_S$ (Quotient aus d_N und d_S) mit d_N als Anzahl nichtsynonymer Substitutionen pro nichtsynonymer Sequenzposition und d_S als Anzahl synonyme Substitutionen pro synonymer Sequenzposition berechnet unter Verwendung des Kimura-2-Parameter-Modell. Die Berechnung wurde beschränkt auf die Lektin-ähnlichen Jacalindomänen der NSPs mit Ausnahme der N-terminal gelegenen Lektin-ähnlichen Jacalindomäne des AtNSP4 (AtNSP4A).

Benennung	<i>Schouwia purpurea</i> (Brassicaceae) NSP	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4B	<i>Cardamine hirsuta</i> (Cardamineae) NSP	<i>Isatis tinctoria</i> (Isatideae) NSP
<i>Schouwia purpurea</i> (Brassicaceae) NSP							
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP1	0,464						
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP2	0,491	0,379					
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP3	0,490	0,293	0,350				
<i>Arabidopsis thaliana</i> (Camelineae) AtNSP4B	0,441	0,247	0,342	0,422			
<i>Cardamine hirsuta</i> (Cardamineae) NSP	0,432	0,310	0,365	0,393	0,359		
<i>Isatis tinctoria</i> (Isatideae) NSP	0,802	0,495	0,566	0,580	0,557	0,624	

7.6 Glucosinolatprofil in *T. arvense*

Tabelle S 9: Glucosinolatgehalt in Organen von *T. arvense*. Der Glucosinolatgehalt für Glucosinolate mit folgenden Seitenketten: Allyl-, Benzyl-, Indol-3-yl-methyl- (I3M), 1-Methoxy-indol-3-ylmethyl- (1MOI3M), 4-Methoxy-indol-3-yl-methyl- (4MOI3M), 4-Hydroxy-indol-3-yl-methylglucosinolat (4OHI3M) und 3-Methylsulfinylpropylglucosinolat (3MSOP) wurde durch HPLC der Desulfoglucosinolate bestimmt. Der Gesamtglucosinolatgehalt je Gewebe ist als „Total“ angegeben. Aufgeführt sind drei unabhängige Einzelbestimmungen (A, B, C). Angegeben ist der Glucosinolatgehalt aufgeführter Organe aus *T. arvense* in μmol pro 100 mg Frischgewicht. Pro Einzelbestimmung gingen je nach Gewebe Material aus ein bis sechs Individuen ein.

A)

Gls	Keimling	Wurzel	Stängel	Blatt	Blüte	Schote	Samen
3MSOP	-	-	-	-	-	0,02	-
Allyl	2,83	0,25	2,56	6,75	19,84	22,28	23,17
Benzyl	0,07	0,08	0,02	0,01	0,20	0,01	0,08
4OHI3M	-	-	< 0,01	-	-	< 0,01	-
I3M	0,01	< 0,01	-	-	-	-	-
4MOI3M	0,08	0,11	0,03	< 0,01	< 0,01	-	-
1MOI3M	0,13	0,14	-	-	-	-	-
Total _A	3,11	0,58	2,62	6,76	20,04	22,31	23,26

B)

3MSOP	< 0,01	-	-	-	-	-	-
Allyl	3,47	0,27	4,07	3,56	8,05	13,63	22,05
Benzyl	0,02	0,06	0,01	0,01	0,01	0,01	0,07
4OHI3M	< 0,01	-	0,01	-	< 0,01	< 0,01	-
I3M	< 0,01	< 0,01	-	-	-	-	< 0,01
4MOI3M	0,02	0,18	0,01	< 0,01	-	-	-
1MOI3M	0,11	0,11	-	-	-	-	-
Total _B	3,63	0,62	4,10	3,57	8,07	13,64	22,13

C)

3MSOP	-	-	-	0,01	-	0,10	-
Allyl	2,05	0,26	0,52	4,30	13,87	14,87	33,75
Benzyl	0,09	0,70	0,03	0,01	0,09	0,05	0,10
4OHI3M	-	-	-	-	-	-	-
I3M	0,01	0,02	-	-	-	-	-
4MOI3M	0,04	0,26	< 0,01	< 0,01	-	-	-
1MOI3M	0,20	0,90	-	-	-	-	-
Total _C	2,39	2,16	0,56	4,32	13,96	15,02	33,85

7.7 Primer

Tabelle S 10: Auflistung verwendeter Primer. Tabellarisch dargestellt sind alle im Rahmen der vorliegenden Arbeit verwendeten Primer unter Angabe ihrer Nukleotidsequenz beginnend mit ihrem 5'Ende sowie der Spezies auf deren cDNA sie zum Einsatz kamen. Die Symbole für degenerierte Basen folgen der Nomenklatur und lauten: S = G + C, Y = C + T, R = A + G, D = A + T + G, W = A + T, K = T + G, M = A + C. Ap = *Alliaria petiolata*, Bv = *Barbarea vulgaris*, Ch = *Cardamine hirsuta*, Ci = *Cardamine impatiens*, C. spinosa = *Cleome spinosa*, Da = *Draba aurea*, Dc = *Draba cinerea*, Dl = *Draba lanceolata*, Ia = *Iberis amara*, It = *Isatis tinctoria*, M. oleifera = *Moringa oleifera*, Sp = *Schouwia purpurea*, Ta = *Thlaspi arvense*, ESP = Epithio-spezififizierendes Protein, TFP = Thiocyanat-formendes Protein, NSP = Nitril-spezififizierendes Protein

Benennung	Sequenz	Spezies / Proteintyp	Verwendung
<u>Degenerierte Primer</u>			
P1 JCK1	TCG ACA AAS ACC TYT ACG TCT TTG ACT TCA A	alle	DOP-PCR
P2 JCK2	TCC CAC ACC ART GTC TCR GTG TCC AA	alle	DOP-PCR
P3 JCK3	GAT GAW AAC CAT GTG TAT RTA TT	alle (ohne Ergebnis)	DOP-PCR
P4 JCK4	GCA AAC CCR YAA ACC ACC CAA A	alle (ohne Ergebnis)	DOP-PCR
P5 JCK5	TCT ATD YCT WYG GWG GCC GDG AC	alle	DOP-PCR
P6 JCK6	TCA TTG STY KGM GMD TTM CCR CCA	alle	DOP-PCR
<u>Allgemeine Primer für 5'RACE- und 3'RACE-PCR</u>			
P7 JCK Anchor d(T)18	GGC CAC GCG TCG ACT AGT ACT TTT TTT TTT TTT TTT TT	alle	5' / 3'RACE-PCR
P8 JCK Anchor	GGC CAC GCG TCG ACT AGT AC	alle	5' / 3'RACE-PCR
<u>Genspezifische Primer für die Isolierung von cDNAs</u>			
P9 JCK GSP 5A Klon 3	CCA AGT CTG AGT GTT GAA GTC	ChESP	5'RACE-PCR (nested)
P10 JCK GSP 5B Klon 3	GGG GCT TCT CCG TTG GCC GGA GCG	ChESP	5'RACE-PCR
P11 JCK GSP 3A Klon 3	GAG GGT TAT GTG TTG GAC AC	ChESP	3'RACE-PCR (nested)
P12 JCK GSP 3B Klon 3	CCC GGG ACG TTG AAC AAT GAG GG	ChESP	3'RACE-PCR
P13 FL USER Klon 3 fw	GGC TTA AUA TGG CTC CAA CTT TTC AAG GCG	ChESP	Expressionskonstrukt
P14 FL USER Klon 3 rev	GGT TTA AUT TAA GCT GAG TTA ACA GCG T	ChESP	Expressionskonstrukt
P15 FL 3 mit UTR 2 F	GAC ACA AAC CTT TGT AGT TTC TAT TG	ChESP	Full-length Klonierung
P16 FL 3 mit UTR 2 R	CAT TTT ATT ATA TAA TAC CAC GTA CGT ACG	ChESP	Full-length Klonierung
P17 JCK GSP 5A Klon 10	GCC GTT GTA TTT ACG TGA AGC GTC	ChNSP	5'RACE-PCR (nested)
P18 JCK GSP 5B Klon 10	GTC GTG TCA AAC GAG TAG AAG CCG	ChNSP	5'RACE-PCR (nested)
P19 JCK GSP 3A Klon 10	GGG CTT GTG ATG CAT GGC GG	ChNSP	3'RACE-PCR (nested)
P20 JCK GSP 3B Klon 10	GGG ATC ATT GAT GGT AAG AAA GGG C	ChNSP	3'RACE-PCR

P21	GSP 5A Klon 2	CCA AGT CCG GGT CTT GAA GTC	identisch ChNSP	5'RACE-PCR (nested)
P22	GSP 5B Klon 2	CCA GTG GCT GGA GAA ATG G	identisch ChNSP	5'RACE-PCR (nested)
P23	GSP 3A Klon 2	GGA TCA AAC CTC TAC GTC TTT G	identisch ChNSP	3'RACE-PCR (nested)
P24	GSP 3B Klon 2	GGT GTC CGT ATG GTT TCA G	identisch ChNSP	3'RACE-PCR
P25	GSP 5C Klon 2 / 10	GCC GTT CCG TAC CAC CCA CTC C	ChNSP	5'RACE-PCR
P26	GSP 5Z Klon 10	GTG GCT GGA GAA ATG GAC C	ChNSP	5'RACE-PCR (nested)
P27	GSP 5Y Klon 10	CGA GGT CAA AGA CGT AAA GG	ChNSP	5'RACE-PCR (nested)
P28	USER Ch10 fw neu!	GGC TTA AUA TGG CCC AAA AGC TGG AAG	ChNSP	Expressionskonstrukt
P29	USER Ch10 rev neu!	GGT TTA AUC TAA TGG ATT CCA TAA AAG AAG	ChNSP	Expressionskonstrukt
P30	FL Ch10 fw	GGA ATC ACA CAA GTC ATA ACT TTG C	ChNSP	<i>Full-length</i> Klonierung
P31	FL Ch10 rev	CCA CGC TTA TCG CAT GGT ATC ACG	ChNSP	<i>Full-length</i> Klonierung
P32	GSP 5A Klon i	GAG TCT TGA AGT CAA AGA CG	ChNSP Fragment	5'RACE-PCR (nested)
P33	GSP 5B Klon i	CCG TTT GGA GGA GCA AGC GAC	ChNSP Fragment	5'RACE-PCR (nested)
P34	GSP 5C Klon i	GGT CCT TTG TTG ACA TGT GTT ATG	ChNSP Fragment	5'RACE-PCR (nested)
P35	GSP 5D Klon i	GGT GGA GTG AGT ATT AGT ACG	ChNSP Fragment	5'RACE-PCR
P36	GSP 3A Klon i	GGG ATA TTG ATA GGT GGG GC	ChNSP Fragment	3'RACE-PCR (nested)
P37	GSP 3B Klon i	GGA CGA AAG GGC TCA TTT GGG	ChNSP Fragment	3'RACE-PCR (nested)
P38	GSP 3C Klon i	GGA TTT AGT GGG AAG GAA ATG G	ChNSP Fragment	3'RACE-PCR
P39	GSP 5A Klon 25	GAG AGT TGA AGT CAA AGA CG	2tes ApTFP ?	5'RACE-PCR (nested)
P40	GSP 5B Klon 25	CCT TTG GCC GGA GCT TTT GAC C	2tes ApTFP ?	5'RACE-PCR (nested)
P41	5RACE GSP1 ApKlon25	TCG GGT CCT CCT TCA TCC AG	2tes ApTFP ?	5'RACE-PCR
P42	5RACE GSP2 ApKlon25	AGC TTT TGA CCA ACT GTG AGA GTT G	2tes ApTFP ?	5'RACE-PCR (nested)
P43	5RACE GSP3 ApKlon25	TTG TCT CCG ATC ACG GCT ATG C	2tes ApTFP ?	5'RACE-PCR (nested)
P44	GSP 3A Klon 25	CCC GTG AGG CTT TCG TGT TGG	2tes ApTFP ?	3'RACE-PCR (nested)
P45	GSP 3B Klon 25	GCA CAT CAA GGT TCG GGA ACG	2tes ApTFP ?	3'RACE-PCR (nested)
P46	3RACE GSP1 ApKlon25	ATC CTA ATA TAC CAC CCA AAC G	2tes ApTFP ?	3'RACE-PCR
P47	Ap25_USER_fw	TAG GCT TAA UAT GGC TCC GAT TTT GC	2tes ApTFP ?	Expressionskonstrukt
P48	Ap25_USER_rev	TTG GTT TAA UTT AGG CGG AAT TGA CTG C	2tes ApTFP ?	Expressionskonstrukt
P49	GSP 5A Klon 18	GAG TGT TGA AGT CAA AGA GG	ApTFP	5'RACE-PCR (nested)
P50	GSP 5B Klon 18	CCC TTG GCC GGA GCG ATT GAC C	ApTFP	5'RACE-PCR (nested)
P51	5RACE GSP1 ApKlon18	TCG GGT CCT CCC TCT TCA TC	ApTFP	5'RACE-PCR
P52	GSP 3A Klon 18	CCC GTG AGG GTT TCG CGT TGG	ApTFP	3'RACE-PCR (nested)
P53	GSP 3B Klon 18	GCA CAT CAA GGC CCG GGG ACC	ApTFP	3'RACE-PCR (nested)

P54	3 RACE GSP1 ApKlon18	ATC CTA ATA TAC CAA CAC TC	ApTFP	3'RACE-PCR
P55	3 UTR Alliaria pet 18 rev	ATA CTC ACG TAC ACG AAA CAC G	ApTFP	<i>Full-length</i> Klonierung
P56	5 UTR Alliaria pet 18 fw	Identisch mit P125	ApTFP	<i>Full-length</i> Klonierung
P57	USER Primer Alliaria 18	Identisch P122, P123	ApTFP	Expressionskonstrukt
P58	GSP 5A Klon 5	CGG GTC TTG AAG TCA AAG ACG	ApNSP Fragment	5'RACE-PCR (nested)
P59	GSP 5B Klon 5	CCT TGG CTG GAG AAA TGG ACC	ApNSP Fragment	5'RACE-PCR (nested)
P60	5 RACE GSP1 Ap Klon5	TTG TAT TTG CGG GAA GCG TCT CG	ApNSP Fragment	5'RACE-PCR
P61	GSP 5C Klon 5	TAA GCA AGA GAG GTG AGG AAC G	ApNSP Fragment	5'RACE-PCR (nested)
P62	GSP 3A Klon 5	CGA TGG GAC TTT TGC GTT GG	ApNSP Fragment	3'RACE-PCR (nested)
P63	GSP 3B Klon 5	CCA CAA GCT CAC GTG GGT CCG	ApNSP Fragment	3'RACE-PCR
P64	JCK GSP 5A Klon Da2 / DI3	CCA CCA TGC GGA CGC CTA AGC	DaESP / DIESP	5'RACE-PCR (nested)
P65	JCK GSP 5B Klon Da2 / DI3	CGG CCT CCA AAG ATA TAG AGC	DaESP / DIESP	5'RACE-PCR (nested)
P66	JCK GSP 5C Klon Da2 / DI3	CGA GTA AAA GTC TTC GAA GTC C	DaESP / DIESP	5'RACE-PCR
P67	JCK GSP 3A Klon Da2 / DI3	CGA AGC CAG ACC CAA ATG CCC	DaESP / DIESP	3'RACE-PCR (nested)
P68	JCK GSP 3B Klon Da2 / DI3	CCC GGC TTC TGG AAA ATG GAC	DaESP / DIESP	3'RACE-PCR
P69	FL Da2 / DI3 USER fw neu	GGC TTA AUA TGG CTC CGA GTT TAC AAG G	DaESP / DIESP	Expressionskonstrukt
P70	FL Da2 / DI3 USER rev neu	GGT TTA AUT TAC GCG GAA TTG ACT GCA TAG AAG	DaESP / DIESP	Expressionskonstrukt
P71	FL Da2 / DI3 mit UTR fw	GGC GAG AAA GAT TCA ATA TTG C	DaESP / DIESP	<i>Full-length</i> Klonierung
P72	FL Da2 / DI3 mit UTR rev	CTT TAT TCC AAT GAC AAC ATA TC	DaESP / DIESP	<i>Full-length</i> Klonierung
P73	5Z Klon Ci1	Identisch mit P9	CiESP	5'RACE-PCR (nested)
P74	JCK GSP 5A Klon Ci1	GAG TAA AAG TCG CAG AAG TGG C	CiESP	5'RACE-PCR (nested)
P75	JCK GSP 5B Klon Ci1	GGT CCT CCC GCC TCA TCG	CiESP	5'RACE-PCR
P76	JCK GSP 3A Klon Ci1	GCG TGT TTG CGC ATG CCG TAG	CiESP	3'RACE-PCR
P77	3B Klon Ci	Identisch mit P12	CiESP	3'RACE-PCR (nested)
P78	FL Ci1 mit UTR fw	CCT TTG TAG TTT CTA TTG CG	CiESP	<i>Full-length</i> Klonierung
P79	FL Ci1 mit UTR rev	CGT ACA TAC TTT GAA AAG TAT TCG	CiESP	<i>Full-length</i> Klonierung
P80	USER Ci1	identisch mit P13, P14	CiESP	Expressionskonstrukt
P81	GSP 5A Klon Dc 47	CCA TTG TAA TTG CGG TCG CC	DcNSP Fragment	5'RACE-PCR (nested)
P82	GSP 5B Klon Dc 47	CGT AGA GTG TAG TGC CAA TGG C	DcNSP Fragment	5'RACE-PCR (nested)
P83	GSP 5C Klon Dc 47	GCA AGC GAG GTC AGG AAC GTC GC	DcNSP Fragment	5'RACE-PCR
P84	GSP 3A Klon Dc 47	CGG AGG TGA GAT AGA TAG GG	DcNSP Fragment	3'RACE-PCR (nested)
P85	GSP 3B Klon Dc 47	GGG CGA GGA GCG TGT TTG CG	DcNSP Fragment	3'RACE-PCR (nested)
P86	GSP 3C Klon Dc 47	GAT CCT GTT GAG AGT AAG TGG	DcNSP Fragment	3'RACE-PCR

P87	GSP 5A Klon Dc 10	GCT TAG TTC CCA CGG CAA CC	DcESP Fragment	5'RACE-PCR (nested)
P88	GSP 5B Klon Dc 10	CGA ACT GTT TAT TCT CAT CGC G	DcESP Fragment	5'RACE-PCR (nested)
P89	GSP 5C Klon Dc 10	CCT GTT CAC CGT ATC GTA CG	DcESP Fragment	5'RACE-PCR (nested)
P90	GSP 5D Klon Dc 10	CCA TTG AGT GGT AAG TGC GAG C	DcESP Fragment	5'RACE-PCR
P91	GSP 3A Klon Dc 10	CCT GGC CGG ACC CAG ACG G	DcESP Fragment	3'RACE-PCR (nested)
P92	GSP 3B Klon Dc 10	CGT GTT TGC GCA TGC AGC AGT AGG	DcESP Fragment	3'RACE-PCR (nested)
P93	GSP 3C Klon Dc 10	GGA GAG AAA CCT TCT CCG	DcESP Fragment	3'RACE-PCR
P94	GSP 5A Klon Sp 57	GGT GAG GAA CGT CTC CTG TGG C	SpNSP	5'RACE-PCR (nested)
P95	GSP 5B Klon Sp 57	CCA ATA GAC ACC ATG CGG ACG C	SpNSP	5'RACE-PCR (nested)
P96	GSP 5C Klon Sp 57	CTT GAA GCG TCA CGG CCT CC	SpNSP	5'RACE-PCR
P97	GSP 3A Klon Sp 57	GCG ATG GAT CCA CAA GCT CAC G	SpNSP	3'RACE-PCR (nested)
P98	GSP 3B Klon Sp 57	GGT AAA CAC ATT GTG ATA TTT GG	SpNSP	3'RACE-PCR (nested)
P99	GSP 3C Klon Sp 57	GCC AGG AGT GTT TTC GCT AGT GC	SpNSP	3'RACE-PCR
P100	FL SpK 57 mit UTR fw	GCA CAA AGC ATT ACT TTC TTC TAA GG	SpNSP	<i>Full-length</i> Klonierung
P101	FL SpK57 mit UTR rev	GAA ACA AAC ATG CAG AGA CTG	SpNSP	<i>Full-length</i> Klonierung
P102	SpK 57 USER fw	GGC TTA AUA TGG CCC AAA AGC TGG AAG C	SpNSP	Expressionskonstrukt
P103	SpK 57 USER rev	GGT TTA AUT TAA GCA GAT TCA ATC CCA TAA	SpNSP	Expressionskonstrukt
P104	GSP 5A Klon Sp 31	GGG TTA CGC GTT GGA GAC TG	identisch SpESP	5'RACE-PCR (nested)
P105	GSP 5B Klon Sp 31	TGG TCC GGG GAA ACT GAC C	identisch SpESP	5'RACE-PCR (nested)
P106	GSP 5C Klon Sp 31	GAC CTG GCC GGA CCC AGA CGG AC	identisch SpESP	5'RACE-PCR
P107	GSP 3A Klon Sp 31	CGT CGT TTT ACA ACG TCG	identisch SpESP	3'RACE-PCR (nested)
P108	GSP 3B Klon Sp 31	CGC CCT ATA GTG AGT CGT ATT AC	identisch SpESP	3'RACE-PCR (nested)
P109	GSP 3C Klon Sp 31	CAT GCG ACG TCG GGC CC	identisch SpESP	3'RACE-PCR
P110	GSP 3A Klon Sp 38	GGG TTT TGC ATT GGA CAC CG	SpESP	5'RACE-PCR (nested)
P111	GSP 3B Klon Sp 38	CGG TCC AGG GAC GTT GAC C	SpESP	5'RACE-PCR (nested)
P112	GSP 3C Klon Sp 38	GAT CTG GCC GGA CCC AAA CGG TC	SpESP	5'RACE-PCR
P113	GSP 5A Klon Sp 38	CCA CAG ACA CCA TGC GGG TGC	SpESP	3'RACE-PCR (nested)
P114	GSP 5B Klon Sp 38	GCG GCC TCC GAA GAG ATA GAG C	SpESP	3'RACE-PCR (nested)
P115	GSP 5C Klon Sp 38	CGA ACG AGT AAA ACT CGT CG	SpESP	3'RACE-PCR
P116	FL SpK 38 mit UTR fw	GAC GCA ACA AAT AGC AAA ATT C	SpESP	<i>Full-length</i> Klonierung
P117	FL SpK 38 mit UTR rev	AAC ACA TAT TTA TTT TCC GAC G	SpESP	<i>Full-length</i> Klonierung
P118	SpK 38 USER fw	GGC TTA AUA TGG CTT CGA CAT TGC AAG GC	SpESP	Expressionskonstrukt
P119	SpK 38 USER rev	GGT TTA AUT TAC GCA GAA TTG ACA GCG TAG	SpESP	Expressionskonstrukt

P120	Thlaspi 5RACE GSP1	AAG ACT CTG GTC TTT GGA ACG	TaTFP	5'RACE-PCR
P121	Thlaspi 5RACE GSP2	ATT GAC CAA GTG TGG GTG TTG	TaTFP	5'RACE-PCR (nested)
P122	Thlaspi USER 2 fw	TAG GCT TAA UAT GGC TCG GAC TTT GC	TaTFP	Expressionskonstrukt
P123	Thlaspi USER rev	TTG GTT TAA UTT ACG TGG AGG AAT TGA CTG C	TaTFP	Expressionskonstrukt
P124	3 UTR Thlaspi rev	ATA CTC ACA CGT ACA TAA AAC ACA C	TaTFP	<i>Full-length</i> Klonierung
P125	5 UTR Thlaspi-fw	AAG AGT TGA ATA TTG CAG CTA TGG	TaTFP	<i>Full-length</i> Klonierung
P126	GSP I.t.o2 3A:	GGC ACA TTT GGG TCC CGG GAC C	ItESP	3'RACE-PCR (nested)
P127	GSP I.t.o2 3B:	GGA GGT GAA ACA TGG CCG GAT CC	ItESP	3'RACE-PCR (nested)
P128	GSP I.t.o2 3C:	CGT GTT CGC GCA TGC CGT TGT AGG	ItESP	3'RACE-PCR
P129	5' Primer I.t.o2	Identisch mit P113, P114, P115	ItESP	5'RACE-PCR
P130	USER I.t.o Klon 2 fw	GGC TTA AUA TGG CTC CGA GTG TGC AAG GC	ItESP	Expressionskonstrukt
P131	USER I.t.o Klon2 rev	GGT TTA AUT TAC GCG GAA TTG ACT GCA TAG	ItESP	Expressionskonstrukt
P132	FL + UTR I.t.o Klon2 fw	ACA AAT AAG TAA AGA GAA AGT GAG	ItESP	<i>Full-length</i> Klonierung
P133	FL + UTR I.t.o Klon2 rev	GCA ACA AGA GCA ATC TCC AAA CCA C	ItESP	<i>Full-length</i> Klonierung
P134	GSP I.t.*o4 3A:	CGA GCC CAC GAG GGT GGT GTG C	ItNSP Fragment	3'RACE-PCR (nested)
P135	GSP I.t.*o4 3B:	GGA CTG AGG AGA GTG AGG	ItNSP Fragment	3'RACE-PCR (nested)
P136	GSP I.t.*o4 3C:	GAT ACA GAG ACA TTA GTG TGG	ItNSP Fragment	3'RACE-PCR
P137	GSP I.t.*o4 5A:	CCA CTC GTT AGT CAG AGT CTC G	ItNSP Fragment	5'RACE-PCR (nested)
P138	GSP I.t.*o4 5B:	CCT GGA AGT CCT TCC TCC ACT GG	ItNSP Fragment	5'RACE-PCR (nested)
P139	GSP I.t.*o4 5C:	GCC GTA CAA GAG TGA TAG C	ItNSP Fragment	5'RACE-PCR
P140	GSP I.t.u2 3A:	GGA TCC ACA AGC TCA TGT GGG TCC	ItNSP	3'RACE-PCR (nested)
P141	GSP I.t.u2 3B:	GCC TTG TGC TAG GAG CGT TTT CG	ItNSP	3'RACE-PCR (nested)
P142	GSP I.t.u2 3C:	GGA CAC AAG TGG AAA CAT ATG G	ItNSP	3'RACE-PCR
P143	GSP I.t.u2 5A:	GGT GGG GAA TGT CTC CGG TGG C	ItNSP	5'RACE-PCR (nested)
P144	GSP I.t.u2 5B:	CCA ACT GAG ACC ATA CGA ACA C	ItNSP	5'RACE-PCR (nested)
P145	GSP I.t.u2 5C:	GGT GGG AAG CGT CTC TGC CTA C	ItNSP	5'RACE-PCR
P146	GSP 2te 5A I.t.u(u)	GGA CTG TAG AAC ATC GCT CGA ACG	ItNSP	5'RACE-PCR (nested)
P147	GSP 2te 5B I.t.u(u)	GGA TGG GAA AGA AAC ATA GGC	ItNSP	5'RACE-PCR (nested)
P148	GSP 2te 5C I.t.u(u)	CCA CCT TAG TCC ACT TTC CAA CC	ItNSP	5'RACE-PCR (nested)
P149	USER I.t.u Klon2 fw	GGC TTA AUA TGG CCA AAA AGC TGG AAG CAA AGG	ItNSP	Expressionskonstrukt
P150	USER I.t.u Klon2 rev	GGT TTA AUT TAA GCA GAG TCA ATC CCG	ItNSP	Expressionskonstrukt
P151	FL + UTR I.t.u Klon2 fw	ATA ACT TTC TTC TCA GGT TAT TAA TTC G	ItNSP	<i>Full-length</i> Klonierung
P152	FL + UTR I.t.u Klon2 rev	GCT GAT AAA GGA TCG TAT AGA AGT GG	ItNSP	<i>Full-length</i> Klonierung

P161	5' Primer Mo Klon1	identisch mit P113, P114, P115	<i>M. oleifera</i> ESP?	5'RACE-PCR
P162	3' Primer Mo Klon1	identisch mit P126, P127, P128	<i>M. oleifera</i> ESP?	3'RACE-PCR
P163	USER Mo fw	GGC TTA AUA TGG CTC CGA GTG TGC AAG GC	<i>M. oleifera</i> ESP?	Expressionskonstrukt
P164	USER Mo rew	GGT TTA AUT TAC GCG GAA TTG ACT GCG	<i>M. oleifera</i> ESP?	Expressionskonstrukt
P165	5' Primer CS Klon3	identisch mit P113, P114, P115	<i>C. spinosa</i> ESP?	5'RACE-PCR
P166	3' Primer CS Klon3	identisch mit P126, P127, P128	<i>C. spinosa</i> ESP?	3'RACE-PCR
P167	USER Cs	identisch mit P163, P164	<i>C. spinosa</i> ESP?	Expressionskonstrukt

Primer zur Amplifikation von Sonden für den Southern-Kapillar-Blot

P153	NSP-Sonde fw	CGA GAC GCT TCA CGC CAA TAC AAC GG	Plasmid-DNA AtNSP1
P154	NSP-Sonde rev	ATG AAC ATC ATC TAC TTC ACA TCC G	Plasmid-DNA AtNSP1
P155	ESP-Sonde fw	TGG ACA TTC CTG ACG AAG CTT GAT GAG G	Plasmid-DNA AtESP
P156	ESP-Sonde rev	CAC AGC ATT AGA CTC ATA GTC ATC TTT GCC	Plasmid-DNA AtESP

Primer für die RT-PCR von *T. arvense*

P157	At1g49240 (Actin 8) fw	AGC TGT TCT ATC ACT TTA CGC CAG	<i>T. arvense</i>
P158	At1g49240 (Actin 8) rev	GAT CCC TGC AGC TTC CAT TCC	<i>T. arvense</i>
P159	Thlaspi RT fw	AGA CCA GAG TCT TAG GCA CCC G	<i>T. arvense</i>
P160	Thlaspi RT rev	GGT ATC TTA GGA TCA TTC GCA GT	<i>T. arvense</i>